DOCTOR OF PHILOSOPHY

Assessing the perceived realism of agent crowd behaviour within virtual urban environments using psychophysics

O'Connor, Stuart

Award date:
2016

Awarding institution:
Coventry University

COVENTRY UNIVERSITY

Assessing the Perceived Realism of Agent Crowd Behaviour within Virtual Urban Environments using Psychophysics

by

Stuart O'Connor

A thesis submitted in partial fulfillment for the degree of Doctor of Philosophy

in the
Faculty of Engineering and Computing
Department of Computing

October 2016

Declaration of Authorship

I, STUART O'CONNOR, declare that this thesis titled, 'ASSESSING THE PERCEIVED REALISM OF AGENT CROWD BEHAVIOUR WITHIN VIRTUAL URBAN ENVIRONMENTS USING PSYCHOPHYSICS' and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed: S O'Connor

Date: 07/06/2017

"Crowds, being incapable both of reflection and of reasoning, are devoid of the notion of improbability; and it is to be noted that in a general way it is the most improbable things that are the most striking."

Gustave Le Bon

COVENTRY UNIVERSITY

Abstract

Faculty of Engineering and Computing
Department of Computing

Doctor of Philosophy

by Stuart O'Connor

Inhabited virtual environments feature in a growing number of graphical applications. Simulated crowds are employed for different purposes; ranging from evaluation of evacuation procedures to driving interactable elements in video games. For many applications, it is important that the displayed crowd behaviour is perceptually plausible to the intended viewers. Crowd behaviour is inherently in flux, often depending upon many different variables such as location, situation and crowd composition. Researchers have, for a long time, attempted to understand and reason about crowd behaviour, going back as far as famous psychologists such as Gustave Le Bon and Sigmund Freud who applied theories of mob psychology with varying results. Since then, various other methods have been tried, from artificial intelligence to simple heuristics, for crowd simulation. Even though the research into methods for simulating crowds has a long history, evaluating such simulations has received less attention and, as this thesis will show, increased complexity and high-fidelity recreation of recorded behaviours does not guarantee improvement in the plausibility for a human observer. Actual crowd data is not always perceived more real than simulation, making it difficult to identify gold standards, or a ground truth. This thesis presents new work on the use of psychophysics for perceptual evaluation of crowd simulation in order to develop methods and metrics for tailoring crowd behaviour for target applications. Psychophysics itself is branch of psychology dedicated to studying the relationship between a given stimuli and how it is perceived. A three-stage methodology of analysis, synthesis and perception is employed in which crowd data is gathered from the analysis of real instances of crowd behaviour and then used to synthesise behavioural features for simulation before being perceptually evaluated using psychophysics. Perceptual thresholds are calculated based on the psychometric function and key configurations are identified that appear the most perceptually plausible to human viewers. The method is shown to be useful for the initial application and it is expected that it will be applicable to a wide range of simulation problems in which human perception and acceptance is the ultimate measure of success.

Acknowledgements

I would like to express my sincere gratitude to all of my supervisors, Dr. Christopher Peters, Dr. Fotis Liarokapis, Dr. Chrisina Jayne and Dr. James Shuttleworth, for their continuous support of my Ph.D study and related research, for their patience, motivation, and immense knowledge. Their guidance helped me shape my research, which led to the writing of this thesis. I could not have imagined having a better range of advisors and mentors for my Ph.D study.

I would also like to thank my family for supporting me emotionally throughout writing this thesis and in my life in general.

Contents

Publications

O'Connor, Stuart, James Shuttleworth, Fotis Liarokapis, and Chrisina Jayne. "Assessing the Perceived Realism of Virtual Crowd Grouping Dynamics." Submitted to Journal, ACM Transactions on Applied Perception (TAP).

O'Connor, Stuart, Fotis Liarokapis, and Chrisina Jayne. "Perceived Realism of Crowd Behaviour with Social Forces." In Information Visualisation (iV), 2015 19th International Conference on, pp. 494-499. IEEE, 2015.

O'Connor, Stuart, Fotis Liarokapis, and Christopher Peters. "An initial study to assess the perceived realism of agent crowd behaviour in a virtual city." Games and Virtual Worlds for Serious Applications (VS-GAMES), 2013 5th International Conference on. IEEE, 2013.

O'Connor, Stuart, Fotis Liarokapis, and Christopher Peters. "A perceptual study into the behaviour of autonomous agents within a virtual urban environment." In World of Wireless, Mobile and Multimedia Networks (WoWMoM), 2013 IEEE 14th International Symposium and Workshops on a, pp. 1-6. IEEE, 2013.

O'Connor, Stuart, Szymon Fialek, Etienne B. Roesch, and Christopher Peters. "Towards procedurally generated perceptually plausible inhabited virtual cities: A psychophysical investigation." Intelligent Agents in Urban Simulations and Smart Cities Workshop, 2012 20th European Conference on Artificial Intelligence (ECAI), 2012.

List of Figures

List of Tables

List of Equations

Abbreviations

2AFC	**2 A**lternate **F**orced **C**hoice
2IFC	**2 I**nterval **F**orced **C**hoice
3D1U	**2 I**nterval **F**orced **C**hoice
AI	**A**rtificial **I**ntelligence
ANOVA	**AN**alysis **O**f **VA**riance
CGI	**C**omputer **G**enerated **I**magery
GLUT	Open**GL U**tility **T**oolkit
HiDAC	**Hi**gh **D**ensity **A**utonomous **C**rowds
JND	**J**ust **N**oticeable **D**ifference
LOD	**L**evel **O**f **D**etail
OpenGL	**Open G**raphics **L**ibrary

Chapter 1

Introduction

Figure 1.1: A Screenshot of the Urban Crowd Simulation developed as part of this research. This specific scene shows various agents navigating through a virtual recreation of Temple Bar in Dublin.

Crowd simulation is the process of simulating large numbers of virtual agents that display distinct collective behaviours, an example of which can be seen in Figure 1.1. The process of simulating huge crowds of intelligent agents in real-time is a challenging task due to numerous different considerations, technical or otherwise (Leggett 2004). Crowd behaviour is typically simulated with artificial intelligence (AI) algorithms. Each of these various algorithms can potentially add a component to the resultant behaviour, which is in many ways the sum of its parts. Examples of variation in behaviour between different algorithms can easily be highlighted; often different pathfinding algorithms can deliver different paths for the same goal or different decision making systems can arrive at a different action to be taken. Advanced algorithms can also be incorporated to add specific behavioural features, such as social forces or even learning capabilities.

Simulated environments, inhabited by virtual crowds of agents, feature in large range of graphical applications. By considering the different purposes that virtual crowds have been employed for in recent years, it is possible to see trend in the adoption of these types technologies. Virtual crowds have played a big role for both simulation and entertainment. Examples of serious applications include simulations for evacuation procedures (Almeida et al. 2013), virtual crowds for specific scenarios, such as bioterrorism (Song et al. 2013), and simulations for the reconstruction of virtual heritage sites (Ulicny & Thalmann 2002). Many of these examples seek to better health and safety in some respect, or inform urban planning and construction, meaning they require a high degree of realism. The high-density autonomous crowds (HiDAC) system pictured in Figure 1.2 shows an example of this type of simulation, in which high density crowds are subjected to dynamically changing virtual environments to induce emergent behaviours such as pushing and line formation.

FIGURE 1.2: The HiDAC System simulating large crowds in a dynamic environments to showcase emergent behaviours (Pelechano et al. 2007).

For entertainment, virtual crowds have been used to create computer generated imagery (CGI) in blockbuster movies and as real-time interactable elements in video games. The games industry in particular develops and utilises virtual crowds in many of its successful franchises. Video games such as The Witcher 3 (CD Projekt RED 2015), Assassins Creed Syndicate (Ubisoft Quebec 2015) and Hitman (IO Interactive 2016), are recent examples of how virtual crowds have become a key element within these ever expanding virtual worlds. Figure 1.3 shows the depth of design for the virtual crowds within the Hitman game. The Witcher 3 won the Game of the Year award among many other accolades, in part due to the authenticity of its inhabited virtual world, which can be seen in Figure 1.4. As a medium, video games are perfectly situated to bring virtual crowds to vast audiences of people, however as time goes on consumer expectation rises, with immersion and perceived realism becoming important factors of success.

FIGURE 1.3: Virtual crowds within the Marrakesh level of Hitman (IO Interactive 2016).

What is common across all these different implementations is a need for realism and perceptual plausibility, with regards to the crowd behaviour. Realism in general and in the context of crowd simulation in particular has been given multiple definitions over time, making it a difficult factor to define for measurement. Indeed, different crowd simulations may even require different types of realism based upon their intended purpose. Taking the previous examples, serious applications will commonly require virtual realism, whereby the simulation is as close to reality as possible (Aschwanden et al. 2008), often by utilising actual crowd data. This is because the simulations will typically be utilised to inform important decisions, which will need to be based on outcomes that have a high degree of precision and predictive ability. Achieving virtual

FIGURE 1.4: The inhabited virtual world of the Witcher 3 (CD Projekt RED 2015).

realism is straightforward by modelling based on the ground truth measured from real-world data. Applications for entertainment, such as video games, however, require a degree of immersion (McMahan 2003) to be successful. This allows the viewer to become hooked to the activity, so that time passes unnoticed and they concentrate fully on the experience. In order to achieve immersion, a degree of perceived realism is important to bring all the different elements together. What a viewer may find perceptually realistic, may not necessarily be what is actually realistic. This means rather than aiming for high-fidelity reproduction of real-world behaviours, in these cases focusing on perceptual plausibility can yield better results (O'Connor et al. 2015).

This all leads us to the inevitable question; how do we ensure that our implemented crowd behaviour is perceptually realistic? This is a question this thesis answers through its application of psychophysics as a perceptual evaluation method.

1.1 Method

There are some core issues when investigating simulating both realistic and perceptual plausible crowd behaviour. Consider actual crowd behaviour; it is typically in a state of flux, with individuals hustling and bustling, back and forth. In traditional psychology, even historically famous psychologists such as Sigmund Freud and Gustav Le Bon have tried to apply theories, such as mob mentality (Freud 1975, Bon 1896), to crowd behaviour with varying results. As such, it is no surprise that achieving realistic and perceptually plausible virtual crowd behaviour is no easy task.

Virtual crowd behaviour is implemented using various AI algorithms (Anderson 2003) in the design of a simulation. Typically, a crowd-based system will include a form of decision making (Luo et al. 2009), pathfinding (Cui & Shi 2011) and steering (Reynolds 1999). These allow it to perceive, think and act, to a limited degree. Typically, these basic algorithms do not produce crowd behaviour that is realistic enough for a simulation to realise its purpose. It is possible however, to implement additional algorithms to augment the simulated crowd behaviour in various ways. For example, by incorporating a social forces model (Helbing & Molnar 1995), certain social dynamics can be synthesised by replicating a number of behavioural forces displayed in human crowds. Some simple examples of these forces can cause agents to be attracted or repelled from one another in given situations.

There exists a large body of research for developing and refining AI algorithms for crowd simulation (Leggett 2004). As this suggests there are two common lines, research towards the development of new algorithms and research for refining existing algorithms. The research for algorithm development often aims to utilise new techniques or models for simulating crowd behaviour (Beltaief et al. 2011, Heïgeas et al. 2010, Kim et al. 2012). This type of research can result in new and interesting behaviours, but often does not consider the perceptual plausibility this behaviour exhibits. Similarly, the research for algorithm refinement can extend existing simulated behaviour to make it more realistic (Lemercier et al. 2012, Mehran et al. 2009), but again this does not necessarily improve the behaviours perceptual plausibility.

Increasing the algorithmic complexity or the number of algorithms does not guarantee an increase in the perceptual plausibility of the crowd behaviour and in some cases, less complex algorithms can produce better results. An example of this is the Boids algorithm (Reynolds 1987), which implements simple steering mechanics to produce a good representation of coordinated animal motion. A simulation based on actual crowd data has virtual realism, in the respect it is close to the ground truth, but it is not guaranteed to have perceived realism as it is possible for parameterised and synthesised behaviour to produce better results (O'Connor et al. 2015).

Since neither the number of algorithms, their complexity or the overall fidelity to ground truth are useful as metrics for perceptual plausibility, another method can be employed.

To judge the perceptual plausibility of simulated crowd behaviour, a method involving perceptual evaluation can be employed to gauge perceived realism. These types of methods are not unheard of for evaluating various graphical aspects, and research exists showing the benefits provided by probing human perception (Ennis et al. 2011, Peters et al. 2008, McDonnell et al. 2007). These types of perceptual methods can involve certain elements of psychology, as probing human perception can be seen in many ways as trying to quantify the psyche.

Psychophysics is a category of psychological experiment methods (Baird & Noma 1978) that attempt to do just that. A psychophysical experiment is employed to probe the perceptions of human viewers, in order to determine specific perceptual threshold values. Psychophysics is often used for auditory and visual stimuli identification; however, experiments have been employed in recent years for other considerations such as graphical comparisons (Melo et al. 2014). The application of psychophysics for virtual crowds, in particular simulated crowd behaviour, has been limited.

Realism is an important aspect for crowd behaviour and it has been demonstrated that different simulations require different qualities of realism. It has been made clear that the efficacy of crowd behaviour simulation must be evaluated against something more sophisticated than adherence to a measured ground truth. In this thesis our framework is presented for employing psychophysical methods to evaluate the perceived realism of behavioural features. Through a three step-methodology, three key behavioural features are subjected to quantitative psychophysical experiments using crowd data derived from real-world locations. The perceptual thresholds are calculated, the optimum configurations examined and the differences between the crowd data and the analysed perceptions are explored.

1.2 Motivation

There are three main motivations for having conducted this research to investigate the psychophysical evaluation of crowd behaviour:

1. The first motivation for conducting this research is to look back and evaluate some of the developments, in this case common behavioural feature for crowd simulation, in order to quantify their success. As has been noted, there is large volume of research towards the development and refinement of algorithms for crowd simulation and their application within certain scenarios. The evaluation of such algorithms has received far less attention, especially when perceptual methods are considered. The field of computer science is very much a field of looking forward, rather than critically looking back. Breaking this trend is a worthwhile endeavour and a worthy motivation for this research.

2. The second motivation is the intriguing concept of comparing the perceptions of what viewers expected, to the reality made evident through real world data. One of the biggest markets for virtual crowds is entertainment, as seen in their application within video games. In this type of a medium, realism is a highly subjective prospect whilst being a crucially required element. By using crowd data as a starting point to create a variety of stimuli for psychophysical analysis, it was possible to compare the perceptions of the viewers to the actual reality. Finding the similarities and differences between reality and perception is not only fascinating, but also holds potential benefits when links can be identified towards ensuring perceptual plausibility.

3. The final motivation is that this research will help to show psychophysics used in the evaluation of specific elements of simulation, potentially leading it to be applied more within the area of computer science as a systematic empirical evaluation method. Researchers such as Rachel McDonnell and Christopher Peters have been an inspiration for this thesis, as they have applied perceptual evaluation towards different elements of simulation. By seeing their success at using human perception for evaluation gave confidence for further delving into psychology and applying found methods towards simulated crowd behaviour. To similarly be an inspiration for future research by highlighting psychophysics as a useful method for repeatable empirical evaluation is an important motivation for this work.

1.3 Scope

A series of perceptual experiments are conducted based on current crowd simulation algorithms. By starting with some of the most apparent behavioural features, such as varying velocity, social forces and grouping dynamics, it has been possible to gain data on some of the most widely utilised methods for simulating virtual crowds. These behavioural features were carefully implemented within a specifically tailored urban crowd simulation. This simulation started as a C++ Open Graphics Library (OpenGL) application, however through a process of refinement it was updated over several iterations and ported to the Unity game engine. Even though this thesis does not focus on developing new methods of crowd simulation, the implementations were constructed with the parameter space and algorithmic variation needed for the multiple configurations required as part of psychophysical evaluation.

The main focus of this thesis is improving the perceptual plausibility of virtual crowd behaviour through the use of comparative psychophysics, in order to evaluate key behavioural features for the purposes of calculating perceptual thresholds and identifying optimum configurations that appear most realistic to the average observer. While crowd data and in turn virtual realism is utilised as base for stimuli generation and experimentation, it should be noted that the focus is on assessing the perceived realism rather than simply making a simulation as close to a measured ground truth as possible. As such, the main application for both the perceptual metrics and the set of methods is towards simulations that are judged through perceptual plausibility and human acceptance. This places the research clearly within the realms of entertainment applications such as video games, however this is by no means the only area, as both the set of methods and perceptual metrics are applicable to serious games and simulations for learning purposes. It should also be noted that while the set of methods is potentially useful various simulations and scenarios, the perceptual metrics obtained are geared towards crowd simulation in busy urbanised areas in cities based in the United States and the United Kingdom. Further limitations are discussed in Section 7.2.

1.4 Contribution

Through a series of comparative psychophysical experiments directed at evaluating some prominent behavioural features, a number of perceptual metrics have been gained with regards to perceived realism. The results will contribute towards the future development of multi-agent systems, particularly those in which perceptual plausibility is the ultimate measure of success. The method itself of identifying a behavioural feature, implementing it into simulation and then psychophysically evaluating it to calculate thresholds and optimum configurations, is applicable to a wide range of scenarios, simulation and algorithms. By showing the initial success of this set of methods through the perceptual metrics, it has contributed an alternative means of evaluating elements of simulation, in particular for crowd behaviour.

Perceptual metrics are presented for three behavioural features commonly associated with virtual crowds, namely varying velocity, social forces and grouping dynamics. These metrics act as guidelines for implementing perceptually plausible crowds, aiding in the development process. With varying velocity metrics, both the velocity range and distribution are considered with perceptual thresholds and the optimum configurations calculated. Similarly, social forces metrics provide insights with regards to thresholds and configurations based on the different weights between the agent forces. Metrics for grouping dynamics highlights the most perceptual realistic group frequency and density for virtual crowds in urban locations. By contributing these perceptual metrics, developers of crowd simulation will be able to utilise both the found optimum configurations and the thresholds for when key behavioural features achieve perceptual plausibility, in order to make the appropriately informed design decisions that will be best for their implementation.

For the purpose of conducting psychophysical evaluation, a virtual crowd implementation that was extensible and could be behaviourally built upon was required to support stimuli creation. By comparing different developments and various crowd simulations, this was achieved through the implementation of several existing algorithms. This research is at its core a contribution in the form of the perceptual evaluation of these robust algorithms through a set of methods; however, by also highlighting how these algorithms can be extended and combined as a basis for successful crowd simulation, see Figure 1.1 in which agents can perceive, think and act, and still be extended with new behaviours and technologies, it provides a virtual crowd pipeline that can be adopted by other developers looking for similar degrees of customisation and flexibility.

1.5 Overview of Chapters

The rest of the thesis has been structured into the following chapters:

- **Chapter 2**, presents an overview of the previous research and related work for the simulation of virtual crowds. This includes an outline of the core AI algorithms and additional algorithms that can alter the crowd behaviour. In particular, the current state of the field is considered and recent implementations of virtual crowds are examined.
- **Chapter 3**, presents an overview of psychophysics, giving a look at its history and noting the different methods that can be applied. Additionally, previous research consisting of relevant perceptually evaluated elements is considered, with a focus on those investigating virtual crowds.
- **Chapter 4**, outlines the adapted three-stage methodology of analysis, synthesis and perception and its application within this thesis. In addition, the advantages and previous successes of this type of approach are considered.
- **Chapter 5**, provides an overview of the development of the urban crowd simulation and its various behavioural features. The implementation choices over the course of the research are considered and the different iterations of development are explored.
- **Chapter 6**, outlines a series of psychophysical experiments to investigate three different behavioural features. Results are examined, perceptual thresholds calculated using the psychometric function and optimum configurations are identified.
- **Chapter 7**, summarises the overall contributions, provides a reflection on the different methods employed and discusses the possible avenues for future work.

Chapter 2

Crowd Simulation

In this chapter, crowd simulation is discussed and the various methods of implementing virtual crowds through the use of artificial intelligence techniques are examined. Research that has implemented virtual crowds and relevant technologies are considered, with the implications of the state-of-the-art related back to this thesis, noting how it furthers the field with respect to perceptual evaluation. In addition, the main algorithms that can be implemented for crowd simulation are highlighted, along with the potential advantages and disadvantages of each approach.

2.1 Artifical Intelligence

Artificial Intelligence, abbreviated simply as 'AI', is a branch of computer science that involves enabling systems or certain aspects of a system to take on certain cognitive characteristics, such as perception or learning, which allows for intelligent decision-making with regards to changing environment or other percieved factors. Crowd simulation is process of modelling large groups of intelligent agents, often within real-time applications, such as serious games for virtual training (Ulicny & Thalmann 2001) or virtual heritage (Ulicny & Thalmann 2002). Crowd simulation is linked to AI through the implementation of its algorithms to tailor crowd behaviour. Virtual crowds have been employed in a number of different arenas, for varying purposes including entertainment, education and serious applications. Often the nature of a virtual crowd depends upon the context in which it is deployed. With the area of crowd simulation and modelling progressing at a rapid rate due to the numerous lines of research that are carried out each and every year, the behaviour of virtual crowds is evolving to some degree with the progression of algorithms and other behaviour altering aspects.

Crowd simulations can vary on their implementation but the three main types are flow-based, entity-based and agent-based approaches (Zhou et al. 2010). Each has its own intrinsic advantages depending on the given needs of a specific simulation, for example crowd size and time scale. Flow-based approaches can simulate highly dense crowds by not considering individuals in the crowd, but by treating it as a continuous flow of fluid. Entity-based crowds model individuals as a set of similar entities, typically in a setup closely relating to a particle system, see Subsection 2.2.4. Agent-based approaches implement agents as autonomous individuals, often capable of reacting to certain events and changes within the virtual environment. These multi-agent systems are the most common type of implementation due to the flexibility granted from the potential to use various AI algorithms to tailor the resultant crowd behaviour. As such, agent-based approaches or multi-agent systems are the type of crowd simulation considered in this research.

The process of simulating virtual crowds comprising of indiviually intelligent agents can be achieved by utilising a number of different approaches; however the most common methods involve employing a collection of collaborating AI models working in tandem in order to animate each agent. Typically, this includes implementing decision making (Luo et al. 2009), pathfinding navigation (Paris et al. 2007) augmented with local steering mechanics (Reynolds 1999) and an agent perception system (Ondřej et al. 2010). Advanced behaviour altering aspects, such as agent learning (Shao & Terzopoulos 2005), social forces models (Helbing & Molnar 1995), physiological factors (Pelechano et al. 2007) and sociological effects (Durupinar et al. 2008) can also be utilised to enhance agent believability. Technologies from the games industry and middleware have also been deployed for simulating crowds within a variety of different scenarios (Szymanezyk et al. 2012, Nareyek 2004), to allow for unique research into specific areas as seen in Figure 2.1. The subsections below outline the core algorithms for a typical crowd simulation.

FIGURE 2.1: An example scenario in an airport terminal using game technologies and middleware to simulate virtual crowds(Szymanezyk et al. 2012).

2.1.1 Decision Making

The decision making system is an important aspect in an AI system (Luo et al. 2009), as it allows for a choice to be made in order for a specific behaviour or action to be selected from a range of possible behaviours or actions. It is up to the decision making system to discern which of these is the most appropriate to choose at that given moment. Often a decision will equate to a new goal or objective which will then passed to another component within the AI system, such as the pathfinding algorithm that would then move the agent towards its new goal. Learning can also be incorporated into the decision making system so that the AI system can adapt to situations that were not originally foreseen. There are different decision making systems that can be implemented into an AI system, each of which has different purposes and advantages that depend upon the systems context. Below is a list of relevant decision making systems:

- **Decision Trees:** A decision tree is essentially just that, a tree containing nodes, connections and leaves (Quinlan 1986). The nodes represent the tests to be conducted, the connections represent the decisions to be made and the leaves represent the resultant actions or behaviours. Generally, each test is of Boolean nature to check a specific variable, however enumeration, numeric range, among other tests are also possible. The main advantage of a decision tree is that it is efficient and quick to process especially if it is balanced, this is mainly because each decision is simply one test.

- **Finite-State Machines:** A finite-state machine has a number of defined states, each of which has actions and behaviours associated with it (Tung & Kleinrock 1996, Gill et al. 1962). Only one state is active at once and it stays as the active state until some condition is satisfied to cause the transition to a new state. Whilst in a state the actions or behaviours associated with it are executed. There are some finite-state machine variations such as the Markov model to add probability to transitions and the hierarchical state machines where a state can transition into its own state machine. Finite-state machines have the advantage of being very fast to process especially if they are expressing simple states and behaviours, however for more complex AI the number of states and transitions can grow exponentially making it become an inefficient method. An example of a finite-state machine can be seen in Figure 2.2.

- **Artificial Neural Networks:** An artificial neural network is designed to mimic the human brain in terms of its architecture, as such it contains multiple artificial neurons that together form a large network, with the axon being represented by the output values (Yegnanarayana 2009). Each neuron has a value associated with it that is calculated as the sum of all its input values, multiplied by its connection weight and added to its bias value. There are three types of layers to the network, the input layer, the processing layers and the output layer. Firstly, the network must be trained by inputting specially prepared data, only then can it be utilised on a live application. The advantages of an artificial neural network lie in its abilities to encapsulate non-linear functions and to adapt to changes in environment or context, however it is very complicated concept and difficult to implement.

There is no one decision-making method that can be called 'best', because each has its advantages and disadvantages, which can be more or less evident depending on the application at hand. That being said for this urban crowd simulation research, neural networks are not be needed due to the limited number of behaviours required for the pedestrian AI. Implementing a finite-state machine is the most appropriate option due to the advantages stemming from the limited number of possible behaviours and the fact that there are different states for pedestrians such as wandering or path following. More detail on core algorithm implementation is discussed in Section 5.3.

FIGURE 2.2: A diagram showing the finite-state machine of a simple enemy in a video game (Bevilacqua 2013).

2.1.2 Pathfinding

Pathfinding algorithms are a key element for any real-time simulation that involves rendering intelligent agents. The main purpose of pathfinding is to move an agent from its current location to the next selected location, which is often passed down from the decision-making process. It might not seem like much of a challenge to begin with but it can be a difficult process to implement (Paris et al. 2007). Firstly, the virtual environment often needs to be split up into nodes and the connections between them, as seen in Figure 2.3. Each of the connections then needs to be defined as walkable or non-walkable, which in some instances can be dynamic as part of shifting environments. This then allows for a graph searching algorithm to be employed to search through the nodes and their connections, between the starting location and the destination, to find a viable path. Only the viable nodes with walkable connections between them are considered for overall path calculation. When simulating huge numbers of agents in a crowd simulation however, this quickly becomes problematic (Botea et al. 2013) and often requires either interruptions before the whole path is complete or an artificial intelligence scheduler.

Below is a list of the key graph searching algorithms that can be implemented for a crowd simulation:

- **Best-First Search Algorithm:** This algorithm does not guarantee to find the shortest path, however its advantage lies in the fact that it is very fast to compute under most circumstances (Korf 1993). It works by calculating a heuristic value for each node based on how close it is to the destination node. It then searches from the starting node straight to the destination node, by moving from the lowest heuristic node to the next lowest heuristic node within its search radius and so on.

- **Dijkstra's Algorithm:** This algorithm guarantees to find the shortest path, however it is slow when compared to the best-first search algorithm (Eklund et al. 1996, Stout 1996). It works by searching in every direction equally without prioritising those that are closer to the destination node. From the starting node it repeatedly searches the closest node that has not already been searched until the destination node is found and thus the path is calculated.

- **A* Algorithm:** This algorithm guarantees to find the shortest path and it searches in the direction of the destination node, but can backtrack if that is unsuccessful unlike the best-first search algorithm (Cui & Shi 2011, Botea et al. 2004). A* is more complicated than the previous two algorithms as it utilises three values for each node. The goal value is the cost from the starting node up to the current node. The heuristic value is the estimated cost from the current node to the destination node. The fitness value is the sum of the previous two values, meaning it is essentially the calculated guess of the cost of a path using this node to get to the destination node. It functions by searching the nodes with the lowest fitness value within the search radius, from the starting node until the destination node is found and then the path can be calculated. In this manner it searches towards the goal but can backtrack if a path proves expensive and it will definitely find a path if one can be found.

A* is, according to the literature, one of the more commonly implemented algorithms especially in simulations and games. This is because its advantages more than make up for its disadvantages in terms of memory usage and processing speed. This is the type of graph searching algorithm that is implemented within the urban crowd simulation.

It makes up the main component of the pathfinding element and allows agents to move from one location to another within the virtual environment. More information on the core algorithms implemented within the urban crowd simulation are presented in Section 5.3.

FIGURE 2.3: An example graph of nodes and their connections for pathfinding in the game Banished (Shining Rock Software 2014).

2.1.3 Steering

Steering mechanics are a common component for agents within simulations and games, as they allow for dynamic local movement changes for actions such as avoidance. In addition, there are specific types of steering such as seek and flee, path following and flocking (Reynolds 1999) that can subtly alter an agent's movement pattern. The type of steering mechanic applied most often depends upon the given scenario and the type of agents being simulated. Most all steering mechanics utilise some kind of perception system (Ondřej et al. 2010), which can be as simple as a defined radius around the given agent. Each agent reads environmental information from its perception system to make local adjustments to its high-level movement routine, depending upon the steering mechanic applied. Below is a list of the most common steering mechanic types:

- **Flocking:** This is the type of steering was first introduced as Boids (Park et al. 2003, Reynolds 1987). It is the combination of three forces utilising a small neighbourhood perception mechanism. The force for separation allows an agent to avoid crowding with other flocking entities. The force for alignment ensures that an agent is steering towards the average heading of other flocking entities. The force for cohesion causes an agent to move towards the average position of the other flocking entities. These forces combined create the flocking steering behaviour which is so named as it appears similar to flocks of birds.

- **Seek and Flee:** This steering mechanic causes an agent to either move towards or away from a specific location, object or agent (Go et al. 2006, Reynolds 1999). A steering force calculated from the agents current velocity and its desired velocity. This steering force is then applied to the direction, towards the target for seek or away from the target for flee.

- **Wandering:** This steering mechanic causes an agent to wander through its environment (Reynolds 1999). This is a type of randomised steering but it has some long term sense of order as the steering direction is derived depending upon the steering direction of the last frame. A wander direction steering force is calculated by taking the current steering force and applying a random offset. The final steering force however is highly constrained to be within a certain limit when compared to the previous steering force, in order to ensure that movement is not too randomised and unrealistic.

- **Arrival:** This steering mechanic causes an agent to move towards a target very similarly to the seek and flee behaviour however the difference is that when the agent is about to arrive at its target, its velocity is reduced on a curve so that by the time it does reach the target location its velocity has become zero (Reynolds 1999). This behaviour can be applied to the previous seek steering force by setting a radius around the target that when entered causes the steering force to be reduced over time.

- **Obstacle Avoidance:** This steering mechanic causes an agent to actively avoid colliding with obstacles or other agents within the environment (Reynolds 1999). The agent projects itself in order to check if it will have any collisions in the near future, if it predicts a collision a lateral steering force is applied that is opposite to the obstacles centre. This lateral steering force combined with a breaking force depending upon the urgency, causes the agent to move and avoid any obstacle within the virtual environment.

- **Path Following:** This steering mechanic causes an agent to follow a defined path whilst keeping within the defined radius of the path (Reynolds 1999). The agents heading is kept on the path by applying a corrective steering force when a future model based on the velocity predicts it will leave the path. Figure 2.4 shows an example of the path following steering mechanic.

The type of steering mechanic implemented is highly dependent upon the scenario and the type of agents being simulated. For the urban crowd simulation in this research, several steering mechanics are implemented. Obstacle avoidance, which is crucial for local adjustments to avoid collisions and crowd path following, which is ideal for pedestrian based scenarios as a combination of the path following mechanic and a separation force for controlling crowd formation. For in-depth implementation details on core algorithms, see Section 5.3.

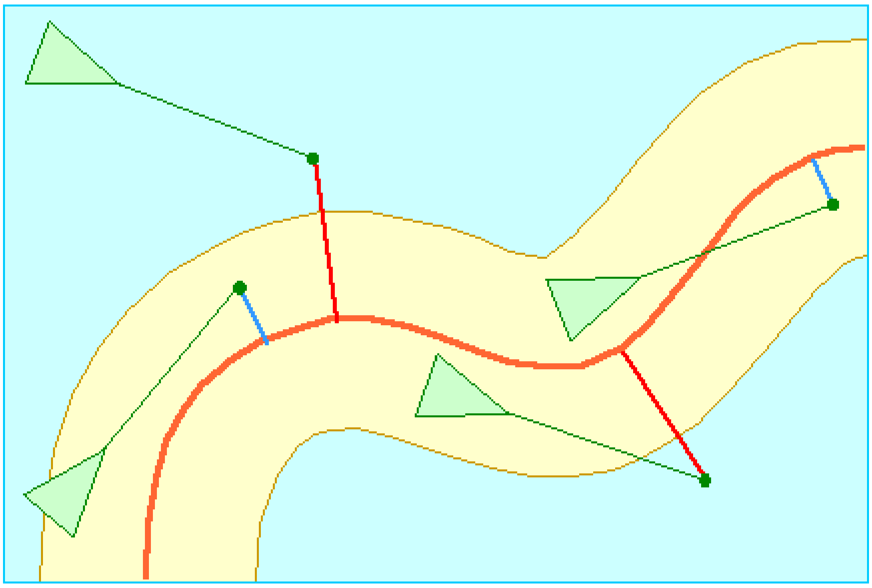

FIGURE 2.4: An example of the path following steering mechanic showing the corrective steering forces of each agent adjusting to the path represented by the red line (Reynolds 1999).

2.2 Crowd Behaviour

A large number of people gathered in a location is the typical definition of a crowd, however the phenomena also encompasses the varying behaviours, which are often the most defining factor of the crowd and not simply the large number of individuals. Figure 2.5 shows a photograph of the famous Shibuya Crossing in Japan and its daily crowds of pedestrians. As can be seen different groups form, some large groups move together while other small groups and individuals move apart. Dynamics within the crowd can cause behaviours to shift, with some pedestrians crossing outside the designated paths, motivating others to do the same.

FIGURE 2.5: A still from a time lapse video of the famous Shibuya Crossing, showcasing its routine large crowds of pedestrians.

In order to simulate a virtual crowd, it is essential that this crowd behaviour can be replicated in some manner for the scene to appear realistic and the crowd behaviour plausible. Figure 2.6 shows a crowd simulation of Shibuya Crossing using crowd behaviour to produce a credible virtual scene (Guy et al. 2012). It is not enough to simply have groups of agents robotically following a predetermined path. Crowd psychology, also known as ‘mob psychology’ was suggested by great theorists such as Gustave Le Bon and Sigmund Freud (Le Bon 2009, Freud 1975). The specifics of their theories varied, however the core concept was that the psychology of the crowd is different from that of the individual, however there is a form of interaction between the individuals and the crowd influencing the resulting behaviours.

For crowd simulation the elements that determine the final crowd behaviour are specific AI algorithms, which often communicate with the core algorithms to some extent. These algorithms will typically try to replicate some form of the human psyche, whether that be social constructs or the internal psychological forces that drive behaviour. This causes a change in the resultant crowd behaviour, whether it be a refinement to core elements such as decision making or adding an entirely new element to the behaviour such as learning capabilities. The subsections below outline some of the relevant behaviour altering algorithms that can be implemented for a typical crowd simulation.

FIGURE 2.6: A crowd simulation replicating the typical crowds of Shibuya Crossing (Guy et al. 2012).

2.2.1 Social Forces Model

Real world crowd behaviour is often driven by social forces; this is especially true when considering human crowds. Often each individual in a crowd will have some predefined social tendancies (Durupinar et al. 2008) that can lead to emergent crowd behaviour. For example, the individual who aggressively pushes through a crowd to reach their destination. As such, a social forces model can be implemented to improve the realism for virtual crowds by simulating these emergent crowd behaviours. There are a number of ways this type of model can be implemented, however one approach, as with the HiDAC system (Pelechano et al. 2007), is for each agent to have a psychological and physiological state. Both states will be checked to gauge an agents overall social state and then the decision making processes and movement will be altered accordingly, leading to emergent crowd behaviours such as impatience, aggressiveness, timidness and even panic in certain scenarios.

Other methods use a similar approach to steering mechanics, where social steering forces are calculated to alter agent movement routines at a local level. These forces are in response to certain stimuli such as avoidance of other agents or attraction to a certain object and so on. It is noted that these types of models can simulate the self-organisation of several collective effects of crowd behaviour very realistically, as with Helbing's original social forces model definition (Helbing & Molnar 1995). Figure 2.7 is an example of this approach, showing the different social forces that can be calculated to refine crowd behaviour. The implementation of social forces in this research is discussed in Section 5.7.

FIGURE 2.7: An example of the social forces that can be calculated in order to add realism to the crowd behaviour (Helbing 2014).

2.2.2 Boids Mechanism

Boids was developed in 1986 by Craig Reynold and is a type of crowd behaviour mainly used for simulating flocks. Boids can be defined as a type of steering mechanic, due to the fact that it adjusts agent's local movement vectors based on three main forces (Reynolds 1987). Each agent has its own local neighbourhood, so that the algorithm only checks other agents within close proximity rather than the whole group. These local adjustments for each individual agent produce an overall crowd behaviour, which appears to give the formation and motion of a flock. As such, Boids is often implemented for simulating animals and swarms rather than humans. An example of Boids used to simulated flocks of birds can be seen in Figure 2.8. The three main forces are separation, alignment and cohesion.

FIGURE 2.8: An example implementation of Reynold's Boids used to simulate vast flocks of birds (de Carpentier 2011). This simulation uses the three forces of Boids with some additional accelerations.

- **Separation Force:** This force adjusts an agents steering vector so that it will avoid local flock-mates.
- **Alignment Force:** This force adjusts an agents steering vector so that it will steer towards the average direction of its local flock-mates movement vectors.
- **Cohesion Force:** This force adjusts an agents steering vector so that it will steer towards the average position of its local flock-mates.

Boids is an important consideration for any research relating to crowd behaviour, due to its massive impact on the field when it was first presented. It is worth noting however, that the behaviour of human crowds is quite different to the behaviour of flocks. As such, the specific implementation may not be viable in this research but the key principles are an important consideration for any crowd-based system and led to the development of future steering behaviours as described in Subsection 2.1.3.

2.2.3 Behavioural Annotation

Behavioural annotation is more of a method rather than a traditional algorithm (Anderson 2003). In this method key information is embedded within features of the virtual environemnt and when interacted with by agents this information causes a shift in behaviour (Peters et al. 2003, Peters & Ennis 2009). This is uncommon as most behaviour altering algorithms such as agent learning (Shao & Terzopoulos 2005), embed the information within the agents themselves. These algorithms however, can often be costly in terms of complexity and processing. The advantage of using the behavioural annotation method is that this cost can be potentially be mitigated to a great extent.

A typical implementation allows agents to perceive the behavioural annotations using their individual synthetic perception. They then alter their own behaviour as their decision-making algorithm takes into account the information within feature they have just seen. Consider a virtual bench, an agent sees the bench and then that information is passed to its decision-making algorithm. This algorithm then decides the agent should rest, so the agent behaviour changes and it takes a seat on the bench. In a typical system this would be internal and all the agents would need an algorithm to check if it is time to rest and then find a seat and so on. By putting this information into the virtual environment, all these checks are mitigated until it is relevant from a simulation standpoint. This urban crowd simulation research includes a behavioural annotation element and implementation is discussed in Subsection 5.4.

It is also possible for additional information such as advanced feature descriptions or local magnitudes to be stored in an annotation, in order to further tailor agent behaviour to their local environment. Authoring (Ulicny & Thalmann 2002) is a similar technique that allows developers to paint crowds and their specific traits using a graphical user interface.

2.2.4 Particle Systems

Particle systems have been implemented for many different purposes in simulation, varying from creating galaxies to modelling fire (Zhou et al. 2006). Figure 2.9 shows an example of a particle system used to produce an effective smoke effect. A particle system typically consists of an emitter from which the particles spawn to live out their life-cycle. During this life-cycle a particle has variables for position and velocity, governed by the laws of physics rather than any specific goal. A sense of randomness is added due to the chaotic nature of the particles and their sensitivity for disparity with respect to initial parameters upon spawning. Additionally, particle systems are utilised to display 'fuzzy' phenomena, meaning a type of fuzzy logic is often incorporated to further the variability and randomness of each particle. Each particles variables are updated each frame of the simulation until its life-cycle is complete and it is removed from the scene. The desired effects such as smoke, are created because masses of particles are spawned and displayed at once, which tends to make the effect look visually impressive.

FIGURE 2.9: A particle system used to create a smoke effect in the video game Modern Warfare 3 (Infinity Ward 2011).

A particle system that would be implemented for crowd simulation is not dissimilar from a particle system for other purposes. This is due to the fact the crowds would follow the same laws and rules as the standard particles (Bouvier et al. 1997). This method of crowd simulation is technically a different type when compared to multi-agent systems (Ferber 1999), however it can be useful in certain instances. In multi-agent systems, each agent is treated as an individual allowing for emergent behaviours from interactions between agents. A particle system however is an entity-based approach and does not allow for this level of depth or intelligence. In essence, the movement of the crowds would possess the characteristics of flowing gas molecules. This has the advantage of being far less computationally heavy compared to multi-agent systems, but does not produce as realistic results in terms of crowd behaviour.

There are a number of methods to implement a particle system for crowds, but typically its just adapting an ordinary particle system for the simulations requirements. This means visually ensuring the particles appear as required, for example representing pedestrians. Other than that the crowds would follow the same physical laws. For a pedestrian scenario the emitter could be situated inside of a building making spawning invisible. The crowds would then exit the building and follow the rules set by the particle mechanism until their life-cycle is complete. They would then promptly enter a building with a particle collector to be removed from the scene. It is possible for the rules of the particle mechanism to be tailored for the specific scenario, in order to make the crowds and their behaviour appear less randomised and unintelligible.

A particle system is not implemented for this research; however, it is an important consideration in terms of how to simulate crowds. Following principles of physical based laws is a possible refinement and the potential for implementing a life-cycle could be incorporated in certain scenarios and for performances purposes. Additionally, by utilising a multi-agent system in this research, it allows for the implementation of behaviour altering algorithms, as detailed in Chapter 5.

2.3 Crowd Simulation Literature Review

By considering the related research for crowd simulation and the implications it has for current research, it is possible to inform implementation and ensure that this thesis fits into the current state-of-the-art for the field. Note, this research consists of two main elements, crowd simulation and perceptual evaluation. This section examines research towards crowd simulation primarily. For details regarding relevant literature for perceptual evaluation, see Section 3.2.

The simulation of virtual crowds has been a popular topic for research within the last decade, as technology has advanced to allow for the real-time graphical representation of hundreds to thousands of intelligent agents simultaneously (Guy et al. 2012). Research time has been devoted to developing specific scenarios for crowd simulation to gain data that could be used for real-world applications such as the optimisation of infrastructures and other avenues of urban planning (Aschwanden et al. 2008). Research has also been devoted to creating more realistic virtual crowds through the further refinement of existing algorithms and the development of new sophisticated algorithms (Guy et al. 2011). Recent crowd based research has had much of its focus on multi-agent based systems (Almeida et al. 2013, Kim et al. 2012) as these offer room for further research and improvement through algorithm design, due to the autonomous and complex nature of the agents. The general overview of these lines of research is to improve the behaviour of the crowds being simulated, within the contexts being placed upon them. This can range from specific scenarios, including certain environments such as an airport (Szymanezyk et al. 2012), to more directed studies regarding single elements of the crowds behaviour, such as collision avoidance (Golas et al. 2014) or agent following (Lemercier et al. 2012).

Emergency and evacuation crowd simulations are often a context of research within specific scenario based studies and serious games. This is due to the advantage that studying actual emergency situations means exposing real life people to danger and possibly hazardous locations, so it is more practice to employ virtual crowds and simulate the behaviour. A recent study into modelling multi-agent systems within this context (Almeida et al. 2013) proposes a possible framework for simulating realistic behaviour within these types of scenarios. This framework suggests utilising a beliefs, desires and intentions technique for the agents, coupled with other features such as social forces and advanced decision making. Another recent study is similar with its context of crowd evacuation for bioterrorism in micro-spatial environments based on virtual geographic environments (Song et al. 2013). By using the theory of virtual geographic environments and bioterrorism assessment, along with crowd simulation techniques, a spatial data environment is developed and utilised to model agent behaviour. The main focus of this study was the development of this framework for implementing a bioterrorism simulation. Another study proposes another framework for using games technology to simulate crowds within an airport scenario (Szymanezyk et al. 2012). This framework utilised PhysX middleware, a spatial database approach, Dijkstra's algorithm for pathfinding and a social forces model. Even though airport specific behaviours were not implemented, the framework itself can be said to be successful in the fact it could produce large crowds at an interactive frame rate. These examples of scenario based crowds suggest a move toward the development of frameworks for solving a specific crowd based problems. In addition, it shows trend with multi-agent systems for complex simulation and the use of algorithms such as social forces for producing generalised crowd behaviour. The implementation of decision making, pathfinding and forms of steering are also prominent and feedback to the core algorithms for crowd simulation as discussed in Section 2.1.

There have been studies toward improving specific elements of virtual crowds to produce more realistic crowd behaviour. One such study considers hybrid long-range collision avoidance for crowd simulations (Golas et al. 2014). This study identifies that local collision avoidance algorithms often ignore agents beyond the local neighbourhood, leading to sharp changes in agent movement. In order to combat this and other issues, a generic algorithm was developed as an extension to current methods to provide a look ahead for achieving long range collision avoidance. Another recent study considers the interactive simulation of dynamic crowd behaviours using general adaptation syndrome theory (Kim et al. 2012). In this study a crowd simulation is developed that incorporates a linearised approximation of the theory of generalised adaptation syndrome for modelling stress response. This is reasonably successful and produces emergent dynamic behaviours, however several simplifying assumptions are made. Another paper deals with realistic following behaviours for crowd simulation (Lemercier et al. 2012). In this paper it is suggested that while many lines of research consider collision avoidance, few have considered grouping. As such a crowd model with following behaviour was developed and calibrated from microscopic analysis of real kinematics data. This model was successful as it matches the needs of a crowd simulation and reproduces the observed phenomena both at macroscopic and microscopic levels. The main aim of these types of specific studies is to enhance virtual crowds in some manner to allow for better realism in terms of behaviour and by extension improved end results. The literature shows that research is currently being carried to produce and improve algorithms, and other elements for simulating virtual crowds. While this type of research has prominence followed closely by employing crowd simulation for scenarios and data acquisition, far less literature considers actually employing in-depth methods for evaluating the resulting behaviour, especially in terms of perceptual plausibility.

For the SIGGRAPH symposium in 2005 a paper was published for research based upon autonomous pedestrians (Shao & Terzopoulos 2005). It detailed about a developed human animation system that utilised perceptual, behavioural and cognitive control components to simulate highly-capable pedestrians in a large scale indoor urban environment. The system proved to be robust and pedestrians carried out individual and group activities successfully. This shows that lines of research have already been devoted to pedestrian simulation and while this specific study covers the intricacies of improving animation through various methods, it highlights that different aspects can be developed to refine crowd behaviour. An important consideration shown within the paper is to form realistic behaviour, the agents need some form of a perception model for interpreting the environment. Artificial perception can be implemented in multiple ways, however a common method as described in this paper is to give each agent a sensing range depicted as a fan shape that originates from the agent in the direction it is facing. The need for agent perception has been discussed in terms of steering mechanics in Subsection 2.1.3. The implementation employs a cell-based system for the virtual environment that can be perceived by the agents to potentially alter behaviour. This is a similar approach to the implementation required for graph searching algorithms used for pathfinding and the method of behavioural annotation, both of which are outlined in Subsection 2.1.2 and 2.2.3 respectively. These are important considerations for the implementation choices made within this thesis and are referred back to in Chapter 5.

A paper published in 2009 by Peters and Ennis outlined a tool for supporting crowds of pedestrian agents in urban environments, named MetroPed (Peters & Ennis 2009). The system was completed and agents can be seen in scenarios as part of the visualisation engine. Future work shows that the system could be improved by implementing a fully automated annotation system for the procedurally generated cities. This shows another line of research towards improving behaviour through development and the implementation of an annotation system for simulating pedestrians. Additionally, it highlights the potential advantages of using a procedural algorithm for generating cities as the virtual environment. These are important consideration for implementation in this thesis and are discussed in more detail in Subsection 5.5. The MetroPed program employs high level graphical interface elements, especially concerning the agent behaviour input. This ensures that even a bespoke program such as MetroPed is straightforward operate and is an ideal extended to this research, as outlined in Section 5.2.

Pelechano et al. presented two papers in 2006 and 2007 respectively. One considered improving the realism of agent movement in high-density crowd simulations (Pelechano & Badler 2006) and the other examined the process of controlling individual agents in high-density crowd simulations (Pelechano et al. 2007). The HiDAC system for simulating crowds of agents is introduced. It uses a combination of social forces, a rule-based model and several unique extensions, in order to produce a variety of emergent behaviours. The overall system uses physiological, psychological and social factors, so that agents can use dynamic information to carry out tasks or actions. The research considered the importance of realism in terms of crowd movement and aims to improve that element. The HiDAC system was tuned to simulate different types of crowds, ranging from extreme panic situations to high-density crowds in calm conditions. In addition, heterogeneous crowds with several different behaviours present at once were also simulated. The research concluded that the HiDAC system could produce more realistic movement through its emergent behaviours, for example by simulating an individual forcing its way through a crowd by pushing others. The system also allowed for more considerate behaviours, which are not typical element present in other algorithms. Overall these papers show there is already a research base for simulating realistic crowds through algorithmic design and help to highlight the great importance of crowd behaviour as the main element for judging crowds. This links very closely to this thesis, which considers both virtual realism and perceived realism in terms of crowd behaviour. This is presented as part of the methodology framework, discussed in Section 4.1.

André Gröschel presented his thesis in 2011 focused towards believable crowd simulation for interactive real-time applications (Gröschel 2011). In this thesis he considers the properties of crowds and their implementation in areas such as video games. Specific interest is given towards making crowds believable and the different algorithms that can be employed. He concludes that entity-based crowds can be favoured because they compute fast, however 'believability' does not depend upon the number of agents but the plausibility of the crowd behaviour. It is suggested that there should be a step away from the technical aspects of crowd simulation and a further look at what viewers perceive as 'social' and 'intelligent' behaviour. This is a highly important conclusion as it again shows that realism in terms of the crowd simulation, in particular the perceived realism or perceptual plausibility of the crowd behaviour is crucial element, worthy of evaluation. In such a technically driven field as computer science comparatively little research time is given to evaluation, showing a need for the type of evaluatory research being presented in this thesis.

It has been shown that lines of research provide advances for crowd simulation, yet there are still some gaps and issues that need to be solved for virtual crowds. One such issue that presents itself is the lack of a unified platform or resources for creating crowd simulations. Most studies will tailor build their crowd simulations from scratch or adapt some form of middleware (Szymanezyk et al. 2012) for their needs. As has been shown with scenario based research, often a framework for how to develop a particular virtual crowd is presented, but due to the time required and the complexity of the task this simulation cannot be fully prototyped and tested (Almeida et al. 2013). This issue hampers researchers as they are often required to tailor build complex software, which takes advanced programming skills and inordinate amounts of time that could potentially be used investigating the main aspects of their research. This issue also affects the development of serious games and if some form of platform or resource database for virtual crowds was available, then many of the assets would not need to be specially created, potentially reducing the games development time. While this thesis does not attempt to solve this issue, it does highlight methods of perceptual evaluation that could be employed to identify core configurations and algorithms to combat this lack of unification. Another issue that presents itself is that crowd simulation encompasses a huge range of specific research areas, from multiple disciplines including computer science, sociology and psychology. As such, there are alot of elements for study and as has been suggested some areas do not get as much focus as others, for example collision avoidance compared to group dynamics (Lemercier et al. 2012). This can lead to some underdeveloped elements for virtual crowds. Current areas of research that could benefit from further exploration include the evaluation of crowd behaviour and looking towards quantum crowds. Quantum crowds have not been discussed in this thesis, however they are being noted as it is an interesting area of crowd simulation research that has potential for further investigation. One line of research applies quantum computing principles to an algorithm for simulating swarms of fish (Zhu & Jiang 2010), which shows the potential of quantum crowd simulation. These are just small sample of the gaps and issues for crowd simulation, but it gives an idea of the challenges present within the field. It also shows how the line of research presented in this thesis fits into the field as a whole, adding to the areas that are not necessarily as researched as others.

A selection of recent research for crowd simulation has been examined and the gaps within the area have been identified, so it is important to consider the possible future trends that could be researched. A highlighted issue was the lack of a unified platform or resource database for virtual crowds and while this is still an area that needs greatly improving, trends show that there is a move towards the use of new technologies that can make the process of creating a crowd simulation far more efficient. One such technology is the games engine Unity, which can be employed to offload a lot of the more troublesome aspects of crowd simulation, such as graphics and animation while providing a stable development environment. These new technologies can also include some aspects of procedural generation for efficient content creation. This move towards new technologies will allow researchers to produce more specialised simulations, more efficiently allowing for further testing of various scenarios and crowd features. This consideration is important in the final implementation of this thesis, and is discussed in Section 5.2. As research progresses, new frameworks will be presented for implementing virtual crowds in serious applications and other contexts. Crowd simulation will continue to be applied to specific scenarios and tested in order to extract data to inform various developments. Research will continue to further develop and enhance AI algorithms that influence crowd behaviour. Virtual crowds will continue to play an essential role within the realms of serious applications such as evacuation procedures, educational applications such as serious games and entertainment applications such as video games and movies.

To summarise with regards to this thesis, various lines of research have been examined showing that the specifics of implementing a crowd simulation can vary depending upon the purpose. However, core AI algorithms make multiple appearances such as decision-making, pathfinding, steering and artificial perception, within multi-agent based systems. Algorithms such as social forces and methods such as behavioural annotation have been implemented to further tailor the resulting crowd behaviour. Additional tools such as game middleware and platforms such as Unity, have been employed to improve various aspects of the simulation such as graphics and interface. All of these elements make up the current state of implementing a crowd simulation and thusly are important considerations that are linked in this researches implementation, discussed in Chapter

5. Finally, the analysed research has shown the importance of behaviour for creating a realistic crowd simulation. Realism both in terms of the simulations accuracy to reality and the perceptual plausibility are fast becoming elements left behind in a technically driven field, showing a need to present methods for evaluating these aspects with respect to the crowd behaviour. That is what this thesis presents and how it is contributing to the gaps present in the current state of the field. This aspect is discussed further in Chapter 4, where the overall methodology of this research is outlined.

2.3.1 Virtual Crowd Implementations

Considering related research is important however, it is also useful to present some instances of recent crowd simulation and the groups focused towards this area. This provides another angle of the field and helps to give a better overall view of the state-of-the-art. Below are some prominent examples of implementations incorporating virtual crowds:

- **EXODUS:** Heralded as the world's most advanced crowd simulation software, buildingEXODUS was developed by Greenwich University fire safety engineering group (University of Greenwich 2012). The package offers crowd simulations for specific scenarios mainly including evacuations and is highly successful at producing realistic crowd behaviour by predicting how individuals will interact with others and the environment. Version 5.0 of buildingEXODUS delves deeper into simulating human psychological aspects in order to improve crowd realism, as such hazards including heat and smoke are taken into consideration for decision making, the results of which are based upon the findings of previous research into the events of those types of situations. Version 6.0 incorporated behaviour associated with lifts for both evacuation and circulation.
- **AnvilNext Engine:** The AnvilNext engine is the current iteration of the Anvil series of specially designed game engines developed by Ubisoft Montreal and mainly utilised in the development Assassin's Creed Franchise of video games. In one of its most recent outing, Assassin's Creed Syndicate (Ubisoft Quebec 2015) the virtual environment of London is populated with highly interactive and realistic crowds. Advanced crowd behaviours are present such as panic and aggression, as well

as some environmental interaction. Virtual crowds are a core gameplay element within these games and the engine will likely be further refined for upcoming games with improved crowd behaviour. While specifics are scarce due to the secretive nature of the games industry, it is highly possible that there will be some new pioneering features in terms of the virtual crowds.

- **T.C Chan Centre:** Based out of the University of Pennsylvania, the T.C Chan Centre is responsible for ongoing research into the topic of building simulation (Liu et al. 2011). Previous research from this centre includes a world trade centre evacuation study and a LOTTE building pedestrian movement study with elevators. Their main area for research at present is into the evacuation procedures of high-rise buildings by employing crowd simulation techniques. This involves the development of a multi-agent communication for evacuation simulation that will give agents multiple roles and individual psychological elements based on studies, in order to simulate virtually realistic crowd behaviour that would be as close to the behaviour displayed during a real evacuation as possible.

- **INCONTROL:** INCONTROL is a company that manufactures simulation software for specific scenarios, including crowd management. Pedestrian Dynamics is the software developed by INCONTROL for simulating large crowds in different infrastructures and in different situations. This software contains state-of-the-art elements including the application of explicit corridor map technology said to originate from the games industry, which produces a unique data structure to represent the continuum walkable spaces of a highly detailed environment. This can be updated in real-time to allow for conditions to be altered during the simulation so that the crowd behaviour can adapted to unanticipated situations such as harsh weather, hazards including smoke and even the structural collapse of buildings. The main application of this software is for health and safety purposes, especially when considering the design of infrastructures and events that gather large crowds. They are currently applying crowd simulation to the infrastructure of airports, more specifically for baggage handling operations (INCONTROL 2016).

These examples highlight some key areas of advancement that are being researched and implemented with regards to crowd simulation. All of these instances aim to improve the realism within crowd simulation and add elements to crowd behaviour. EXODUS brings aspects of human physiological elements to its simulation, whereas the AnvilNext engine incorporates far more interactivity for its crowds. Overall these examples add to the concepts presented in Section 2.3.

2.4 Summary of Chapter 2

In this chapter, the AI algorithms for simulating virtual crowds and influencing the resultant behaviour have been examined. The typical implementation of a crowd simulation has been outlined, with the core algorithms for decision making, pathfinding and steering analysed. The advantages of different algorithms have been noted and related back to the implementation within this research. In addition, crowd behaviour has been discussed and the additional algorithms that can be implemented to refine behaviour in the context of simulation have been explored. This information is important and referred to again in Chapters 4 and 5. The relevant literature relating to crowd simulation is reviewed and the implications linked back to this research. The state of the field is discussed as a whole to show how this thesis augments current research. Finally, recent implementations of virtual crowds have been examined to give a full view of crowd simulation.

In the next chapter, perceptual evaluation is analysed with the history and methods of psychophysics. The early roots of the field are examined through Weber's law and the Fechner-Weber law. The different methods of psychophysics are explored, including constant stimuli, staircase procedures and forced choice. The advantages and disadvantages of these methods are considered and threshold identification through the psychometric function is outlined. Finally, relevant literature employing perceptual evaluation and psychophysics are discussed relating back the implications for this research.

Chapter 3

Perceptual Evaluation

In this chapter, perceptual evaluation is examined with specific focus on a core element of this research; the psychological based model of psychophysics. Similar research that has applied perceptual evaluation as a means to evaluate some element of simulation is considered, along with the potential to extend these methods toward virtual crowd behaviour. Due to the scope of psychophysics in this research and its limited saturation in the computer science field, a timeline of its development and refinement through the years is presented. In addition, the different methods within the area of psychophysics are detailed and a comparison of their advantages and disadvantages in the context of this research is outlined.

3.1 Psychophysics

Psychophysics can be categorised as a type of perceptual evaluation methodology; however, it has its roots planted firmly in psychology. Due to this is has a variety of different methods available, further detail in Subsection 3.1.1, that have some distinct benefits over traditional perceptual evaluation experiments. A main difference is that it allows for the calculation of perceptual thresholds through the psychometric function, explained in Subsection 3.1.2. Given psychophysical methods are being employed in this research and the limited saturation of psychophysics in general computing, it is worthwhile to discuss the early work that created the basis for this branch of perceptual evaluation.

A general definition for psychophysics;

"Psychophysics is commonly defined as the quantitative branch of the study of perception, examining the relations between observed stimuli and responses and the reasons for those relations" (Baird & Noma 1978).

This may be considered a narrow view to some extent as psychophysics has had many varying applications (Roederer 2008), but the definition does highlight the core element of psychophysics. Namely the measurement of the relationship between the stimuli and the participants response. It is possible for the method and context to vary drastically from experiment to experiment, however this core principle remains the same and is the basis for psychophysics.

Psychophysics were initially introduced by Gustav Theodor Fechner in 1860, under a transdisciplinary research program named 'Psychophysik'. The goal of this program was to present a scientific method for studying the relationship between the physical and phenomenal worlds. Fechner's basis for psychophysics was that the body (outer psychophysics) and the mind (inner psychophysics) are simply different views of the same reality. From this he suggested that the processes of the brain are directly reflected in the processes of the mind, which is essentially one of the main goals of modern neuroscience, as it is establishing correlations between neuronal (objective) and perceptual (subjective) events (Fechner 1966). Figure 3.1 shows this relationship between stimuli and perception, using Fechner's concept of inner and outer psychophysics.

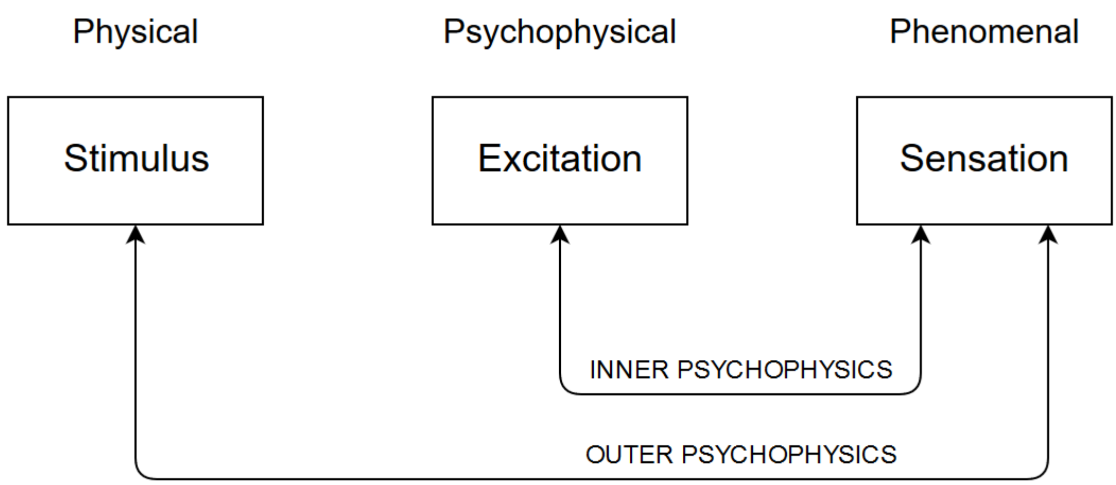

FIGURE 3.1: Fechner's concept of inner and outer psychophysics (Fechner 1966).

Fechner was inspired to his work by German physiologist Ernst Heinrich Weber and his research into the sense of touch and light in the 1830's (Weber 1978). Weber's work with weight perception in 1850 identified the just noticeable difference (JND), which is the minimum amount by which a stimulis intensity must be changed in order to produce a noticeable variation in sensory experience (Ekman 1959). JND, referred to as Weber's Law by Fechner is shown in Equation 3.1.

$$\frac{\Delta I}{I} = k \tag{3.1}$$

Where I represents the original intensity of the stimulus and ΔI represents the difference threshold, which is the minimum amount the stimulus needs to change to be noticeable. k is a constant that signifies the left side of the equation stays constant despite any variation with I.

Weber's law influenced Fechner in the development of his logarithmic Fechner Scale for attempting to describe the relationship between magnitudes of physical stimuli and the perceived intensity of the stimuli. The joint work of both Fechner and Weber became known as the Weber-Fechner Law. Presented in Equation 3.2, this law shows that the perceived intensity is proportional to the logarithm of the stimuli (Portugal & Svaiter 2011).

$$p = kIn\frac{S}{S_0} \tag{3.2}$$

Where p is the perception value, meaning if $p = 0$ then nothing is perceived. k is a constant factor to be determined experimentally and In is the natural logarithm. S represents the stimulus level at that time and S_0 represents the threshold below which the stimulus is not perceived at all.

Charles S. Peirce and Joseph Jastrow who were his student between 1885 and 1886, studied and eventually extended Fechner's work to confirm most of his empirical finding (Peirce & Jastrow 1884). One that was rejected was Fechner's estimated threshold for weight perception, which Peirce and Jastrow found to be too high. Peirce and Jastrow's main contribution was their invention of randomised experiments, which inspired many in the field of psychology and education. They both conducted multiple experiments into the realms of human perception, for stimuli such as weight and light.

In 1942, Hecht, Shlaer and Pirenne defined the absolute threshold for vision in an experiment designed to measure the minimum number of photons that are detectable by the human eye (Hecht et al. 1942). The absolute threshold is the minimum intensity of stimuli needed for a human to perceive it. As this is not always so absolute as in the vision experiment, it can also be defined as the lowest intensity stimuli that is detected at least 50% of the time.

All of the work up to this point by Fechner, Weber and those that helped refine the initial methods, created a base for the field of psychophysics by showing the links between physical stimuli and perception could be successfully probed. After this founding, a great deal of research has employed psychophysical methods for various purposes, however all inevitably draw from these early experiments, as does this research.

3.1.1 Psychophysical Methods

There are various psychophysical methods for presenting stimuli during experimentation (Gescheider 2013). By utilising psychophysics for testing, individual perceptions can be investigated, communicated and shared to others through the link between perceptual experience and physical stimuli. The basis of psychophysics is to use physical stimuli and its intensity values as a reference point for human perception. "The art of psychophysics is to formulate a question that is precise and simple enough to obtain a convincing answer" (Ehrenstein & Ehrenstein 1999). This statement holds true for most psychophysical investigations that often consist simple questions as to whether stimuli can be detected and identified, however other factors can also be considered such as the scale of the stimuli.

Psychophysical methods can be broken down into classical methods and adaptive methods. Classic methods are essentially the processes originally devised by Fechner, however many have been refined and modernised. The classical approach sees data being collected from across the psychometric function, often due to the psychometric threshold being unknown. This provides a shape for the psychometric function but at the cost of efficiency as it is possible the collected data will not be refined to the threshold point. In contrast, adaptive methods, which were developed later through refinement see the sample points tailored to the experiments participants. This means that data is clustered around the threshold point, but at the expense of the overall shape. The psychometric function is discussed in Subsection 3.1.2.

Psychophysics is often focused on the principle of detection thresholds. The two key types of threshold that are prominent in psychophysics are the **absolute threshold** and the JND (Ekman 1959), often just called the **difference threshold**. The absolute threshold is the lowest intensity required for stimuli to become consciously noticeable, which is essentially when the stimuli becomes first detectable to the persons perception. The JND threshold by comparison is not a specific level of intensity but the amount of intensity change required to produce a noticeable difference in the stimuli so that it is detectable by human perception. Weber's law can be used to calculate a constant value from the difference between the initial intensity and the altered intensity, which can be used to predict the amount of intensity change required for a given intensity so it becomes detectable to a persons perception. Weber's law can be seen in Figure 3.1. The JND is also linked to the Fechner-Weber law for calculating the perceived intensity stimuli as seen in Figure 3.2.

Commonly, the identification of units of measurement and scale can be a drawback in psychophysics (Borg 1990). This can be mitigated however as typically the goal is to have metric qualities similar to physics, such as absolute zero and even scale values. This means by using the absolute threshold and the JND, the absolute zero for perception can be determined along with an even scale based on noticeable perceptual incrementation.

Some of the prominent psychophysical methods and their basis for deployment in perceptual experimentation are presented in the following subsections.

3.1.1.1 Method of Adjustment

The method of adjustment is a classical approach that essentially places the stimuli into the hands of the human participant (Ehrenstein & Ehrenstein 1999). In this type of experiment, a participant will be given some form of control over the stimuli and can alter its level of intensity (Farell & Pelli 1999). These types of tests often start from low intensity stimuli and the participant can increase the intensity until it becomes detectable. The reverse is possible whereby it starts at a high intensity and it's reduced by the participant until becoming barely noticeable. It's possible for several runs of adjustments to take place in one experiment including both ascending and descending intensity, in order to improve accuracy by calculating the average thresholds. Using this method is an efficient means for calculating the absolute threshold and the difference threshold as the participant essentially chooses the values themselves.

3.1.1.2 Method of Limits

The classical methods of limits is often utilised for a single stimulus, which has its intensity altered in successive but discrete steps (Ehrenstein & Ehrenstein 1999, Yarnitsky et al. 1995). Each time the intensity is changed the participant's response is recorded, in some cases this can be a simple yes or no response to a question such as "Is this realistic?". Similar to the method of adjustment the stimuli intensity can either start low and be ascended or start high and be descended. It is suggested that if the intensity is ascending or descending, it can cause a slight variation in results so for greater accuracy multiple steps can used for calculating averages. If a single stimulus is utilised, then the absolute threshold can be discovered as the participant will identify when the threshold has been reached.

The method of limits can also be employed with multiple stimuli for comparison, in order to discover the difference threshold. For example, a constant stimulus can be presented next to another identical stimulus that is either increased or decreased in successive but discrete steps until the participant detects a change. The difference between the intensity of the constant and altered stimuli can be used to calculate the difference threshold using Weber's law. As with the stimulus approach, several passes of ascending and descending intensity can be utilised for increased accuracy by calculating the averages.

3.1.1.3 Method of Constant Stimuli

The method of constant stimuli is perhaps the most frequently utilised classical approach as it offers an element of randomness (Ehrenstein & Ehrenstein 1999). This helps to mitigate possible data inaccuracy by the human participant predicting the changes in stimuli intensity (Simpson 1988), which can be possible with the methods of adjustment and limits. Using this method several stimuli intensities, normally between five and nine, are preselected. These selected intensities should cover the thresholds, which are often discovered through analysing key data or using previous experiments, such as those employing the methods of adjustment or limits. The stimuli is then presented in a quasi-random manner, which means that all stimuli will be shown an equal amount of times but in a varying order. The participant is then generally asked if the stimuli is detectable and if it is weaker or stronger than an already presented standard, however depending on the specifics of the given experiment this can be altered (Brunstrom et al. 2008). The final results for each stimuli intensity are then calculated and often plotted on a graph to represent the psychometric function, as discussed in Subsection 3.1.2.

3.1.1.4 Staircase Procedure

The staircase procedure is an adaptive method in which the stimuli intensity is adapted to the individual human participant (Ehrenstein & Ehrenstein 1999). This means that a smaller range of intensity is typically examined, as the majority of data is acquired from around the threshold point. The staircase procedure in itself is an adaptive version of the classic method of limits. Similarly, it generally starts with a high intensity and the participant is asked if they can detect the stimuli. At this level they will be able to, so the intensity will be reduced a step. This continues until the participants can no longer detect the stimuli and instead of ending as it would normally in the method of limits, the experiment reverses, the intensity gets increased a step and the participant is asked the same question. This time the intensity is increased a step each time the participant still cannot detect the stimuli. This continues until they do detect the stimuli, at which point the experiment reverses again. These reversals usually occur between six and nine times in a normal experiment. It is in this manner that data is gathered from around the threshold and the threshold can be calculated as the average intensity at which the participant's answers changed, essentially the transition point.

One issue with the basic staircase procedure that is solved by other methods such as constant stimuli, is the participant predicting the threshold and intensity changes. The constant stimuli method overcomes this with its element of randomness, in essence hindering any participants attempts at prediction. The staircase procedure can be improved in a similar manner, such as with the double staircase-method whereby two staircases are employed. One staircase starts at a high intensity stimuli and proceeds in a descending order, whilst the other starts at a low intensity stimuli and proceeds in an ascending order. The stimuli presented to the participant is then interleaved between the two staircases to essentially disorder the intensities so that they become hard to predict. It is possible to add even more staircases or randomise the interleaving, if an experiment requires it (Gracely et al. 1988, Cornsweet 1962). Utilising a method such that of parameter estimation by sequential testing, accuracy can be further increased as intensity step size can be altered for subsequent trails based on the results from previous trails. There are other variants of the staircase procedure for different purposes, such as the Bayesian model and QUEST (Watson & Pelli 1983). Research has found that the staircase procedure is more efficient and in some cases more accurate when compared to its classical counterpart, the method of limits (Nachmias & Steinman 1965).

3.1.1.5 Magnitude Estimation

Magnitude estimation is an adaptive method that revolves around the participant ordering the stimuli in terms of what they perceive the intensity to be. In a basic experiment a participant will be presented with a series of stimuli with no given order and asked to assign value to each one for what they believe the intensity is (Ehrenstein & Ehrenstein 1999, Poulton 1968). The valuing system may be integers or decimals depending upon the experiment or the participant. Normally the participant can use any range of numbers for assigning values, however it is also possible for a stimulus to be already assigned a value for the participant to use as a reference point. It could be assumed that this flexibility for assigning values would cause severe difference in data between participants, however this is often not the case and the order and spacing between individual participant's values are generally quite similar (Bard et al. 1996). A large range of stimuli intensities are preferable for magnitude estimation experiments. Participants should generally be reminded that it is an estimation experiment if no standard value is used as this can alter the results, and depends upon the experiment. Data from a magnitude estimation experiment can be analysed by using different plotting methods, such as on double-logarithmic coordinates or as the psychometric function, see Subsection 3.1.2. If the data is likely to be similar between participants it can be averaged utilising the geometric mean, but if required the median can be used for some zero values.

3.1.1.6 Forced Choice

Most other psychophysical methods utilise the participants subjective views in order to determine what has been perceived. Forced choice methods on the other hand can offer more objective results by forcing the participant to give an answer on every trial regardless of what has been seen in terms of the stimuli (Ehrenstein & Ehrenstein 1999, Emerson 1986). In a common experiment the stimuli is present in some trials but not in others, so if the participant is unsure they have to guess. As both the correct and incorrect responses are recorded it is possible to estimate the discriminability independent. It is possible that forced choice methods can determine a lower absolute threshold than purely classical methods (Sekuler et al. 2002, Wier et al. 1976), such as the constant stimuli approach, which essentially shows that in a non-forced experiment more stimuli intensity is needed to support a decision. To determine both the absolute and difference thresholds two different tests are often employed, detection tests and discrimination tests respectively. Often the psychometric function, see Subsection 3.1.2, is calculated at 75% for forced choice experiments rather than the usual 50% for other psychophysical methods.

A common method is the two alternate forced choice (2AFC) psychophysical method (Ulrich & Miller 2004). In this procedure two stimuli are presented simultaneously and there is a delay to allow for a response from participants, which is a choice for one of the two stimuli. 2AFC has been utilised for testing the discrimination of motion perception (Gold & Shadlen 2000), a similar area to this research which is based on the motion of virtual crowds to an extent. A variation of the 2AFC is the two interval forced choice (2IFC) psychophysical method (Relkin & Pelli 1987). More choices can be potentially added to these methods for increasing the efficiency (Schlauch & Rose 1990). A forced choice method can be incorporated with other psychophysical methods such as the staircase procedure (Garcıa-Pérez 1998) or the method of constant stimuli. This for example can have the advantages of objectivity and data clustered around the threshold point.

3.1.2 Psychometrics

The accuracy and analysis of data is a highly important aspect in any given experimentation, however perhaps even more so in psychophysics given the subjective nature of human perception. The common approach for many of the psychophysical methods is to plot the psychometric function (Klein 2001, Ehrenstein & Ehrenstein 1999, Wetherill & Levitt 1965), which is a special case of the general linear model. A psychometric function when plotted on a graph shows the relationship between the stimuli intensity and the participants subjective response. Generally a psychometric function is plotted with stimuli intensity on the abscissa and a response, such the number of "yes" responses on the ordinate. Figure 3.2 shows an example of this using another alternative, "probability of correct responses". Often the results will show a sigmoid function resulting in a sigmoidal (S-shaped) curve, which is mainly due to the fact that lowest stimuli intensities are often not detected whereas high intensities are detected the majority of the time, and the intensities around the threshold are detected sometimes but not others. From a plotted psychometric function the absolute threshold can be determined, usually by tracing the 50% response mark on the ordinate and linking it to the stimuli intensity value on the abscissa.

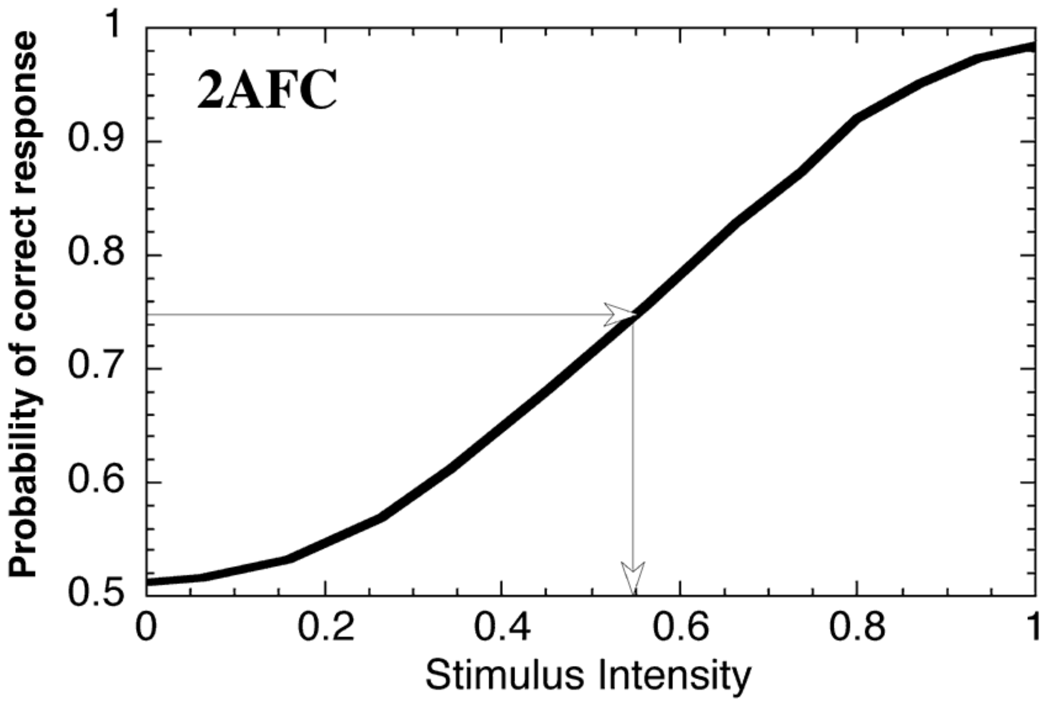

FIGURE 3.2: This is an example psychometric function for the 2AFC method, in which the threshold is taken at the 75% level.

3.2 Perceptual Evaluation Literature Review

In the context of this research, perceptual evaluation or more specifically psychophysics is primarily being used in conjunction with virtual crowd simulation, see Section 4.1 for the overall methodology. The nature of psychophysics being quantification of the human psyche in terms of perceptual subjectiveness, allow it to be linked with the mathematical nature of computer science for experimentation (O'Connor et al. 2013*a*, Brosnan et al. 2005). By examining relevant literature that combines these elements, it is possible to highlight areas of success and the gaps that need to be filled. Perceptual evaluation, psychophysics in particular has comparatively limited saturation in terms crowd simulation research when compared to other areas, partly due to it being a psychology based discipline. There is a breadth of research towards the simulation of crowds, which encompasses multiple avenues of exploration including algorithm design, optimisation and development through simulation, see Section 2.3. Specific work towards using human perception to evaluate virtual crowds is small by comparison, meaning there is room for further investigation to apply these perceptual methods to different elements of crowd simulation or more specifically crowd behaviour. Below some of the prominent related literature is considered and its implications to this research discussed.

Research conducted by Melo et al. consisting of the evaluation of high dynamic range video tone mapping for mobiles devices (Melo et al. 2014), employed perceptual evaluation techniques. The perceptual experiments consisted of ranking seven high dynamic range video sequences that were tone mapped via six tone map operators. The experiment had three independent variables, namely the six tone map operators, the different displays and the scene groups. Sixty participants were split into four groups, two for each display by two for each of the different scenes. Due to possible bias such as preference for brightness, participants were asked "which video sequence is closest to the reference?" rather than "which video sequence do you prefer?". The participants viewed the six tone mapped videos for the scenes they were given and asked to drag thumbnails into empty numbered slots to indicate there ranking via tailored software. Double clicking any of the thumbnails caused the corresponding video to be shown and once all videos were ranked a different scene would be selected until the experiment was completed. The results were averaged and presented for each scene. They were analysed using a mixed design factorial analysis of variance (ANOVA) for two group by two display by the six tone map operators. Kendall's co-efficient of Concordance 'W' was used to give an estimate of agreement amongst participants. The main conclusion highlighted by the results is that while a tone map operator may be the best ranked for one video, it is not the best ranked for all videos. Factors such as the videos attributes must be taken into account, furthermore that the display also has an impact on the tone map operator preference. The outcome of this research shows that perceptual experimentation methods can be successfully used in order to quantify human perception between varying visual material. Although this research did not specifically use psychophysical methods or calculate thresholds, the procedure employed as part of the experiment shares many similarities with psychophysics. Additionally, by using reference material bias can be overcome to a degree which is important for the experiments presented in this thesis in Sections 6.2, 6.3 and 6.4.

Existing research highlights that psychological methods can be successfully applied to certain aspects of crowd simulation. McDonnell et al. has published a series of papers using psychophysical methodologies to assess certain graphical elements and animations of agents, the first of which focused on evaluating the level of detail (LOD) effects regarding the cloth on virtual humans (McDonnell et al. 2006). With current consumer graphics hardware, displaying large crowds of agents with fully deformable clothing is typically not possible in real-time and so often LoD techniques are employed to switch between high and low quality models. Using psychophysical methods, high resolution models were tested against their lower resolution counter parts and an imposter based technique. It was discovered that both high resolution models and imposters could reproduce the fluidity levels of cloth more effectively than low resolution alternatives. By using a forced choice method, it was found that participants judged the low resolution models to have the stiffest cloth, whereas imposters were less stiff and high resolution models were the most fluid. It was determined that the imposters clothing was closer to the ideal and thus more effective when replacing the high resolution models, keeping the virtual crowds as realistic as possible. Further studies identified a perceptual metric for the most appropriate update frequency for the different imposter types. Finally the effectiveness of these insights were perceptually evaluated in a full crowd based system, and results complimented the earlier studies with imposters perceived as the most realistic substitute. An unexpected result led to the insight that the crowd size did not have an effect on the perception of differences between the character types.

A further study extends the previous work, in order to assess the perceptual thresholds for smooth animation in the form of a pose update rate (McDonnell et al. 2007). Measured as poses per second, it is defined as the frequency of updates that display when animating a character. This is an important factor to be assessed, since despite technological advances the utilisation of resources is still an important consideration and the perceptual significance of certain aspects of animation have not yet been explored. From the results of several psychophysical experiments, certain metrics are discovered that can provide good guidance for developers wishing to implement animated characters and crowds. From a movement based study considering the poses with linear velocity, it was determined that for the animation to be perceived as smooth the rate at which the pose is updated needs to be around 40 poses per second. This rate was selected as it is when the probability of acceptance curves become linear, implying that 40 poses per second is the rate at which participants considered all the motion to be smooth 100% of the time. As such increasing the poses per second rate above 40, would not yield any higher results as participants already judged it to be smooth animation. This metric utilised in simulation development would mean that resources are not wasted on a higher pose update rate when it is not required and can be applied elsewhere.

Peters and Ennis conducted an analysis of pedestrians using a corpus of static images and video data picturing some of the areas around Trinity College in Dublin (Peters & Ennis 2009). This research towards modelling groups of plausible virtual pedestrians provides some insights into using perceptual methods to assess crowds. Using MetroPed a specifically developed tool, these analysed scenes could be recreated. However, it was hypothesised that some of the new modelling techniques may affect the viewers perceptions of the scene plausibility and thus these factors should be explored. As a main variable, the group ratios were considered. This consisted seven different ratios showing different numbers of individual pedestrians, duos and trios. An additional 'no group' condition was also used. These scenes were recorded as short 2-3 second video clips to be utilised during the trials. Participants judged the various clips on their plausibility. The researchers hypothesised that the ratios with more individuals and pairs of pedestrians would be more plausible than the other ratios. Results from the perceptual studies indicated that adding groups to crowd-based simulations does increase the plausibility of the scene. This comes with a stipulation however, as the ratios are an important consideration. It was found the ratios that favoured either more individuals or groups of two were the most plausible for viewers, linking back to the researchers' hypothesis. This provides practical insights into the development of virtual crowds, highlighting that perhaps more attention should be given to the behaviour of individuals or pairs, rather than groups of three or more pedestrians.

An additional study by Ennis et al. considers the perceptual effects of scene context and viewpoint for virtual pedestrian crowds (Ennis et al. 2011). One of the perceptual experiments presented participants with various images of character formations, some being real and some being synthetic, and they were asked if they thought a given image was real or not. The results of the experiments suggested that a viewer's ability to distinguish between real and artificial scenes depends heavily on the context of the scene and how the agents adhere to that context. This type of consideration would benefit virtual crowds implemented in mediums such as film, video games and serious games, wherein viewer perception can be an important part of the experience.

This related research has highlighted the use of perceptual and psychophysical methods to evaluate some elements of agents and crowds. It has become apparent that by assessing human perceptions, clear insights into how the crowds are perceived can be identified. Often these perceptual studies will result in additional unexpected discoveries, which indicates how a certain factor might be of more significance than was initially thought. This can lead to finding key configurations and specific aspects that are important for the perceptual realism of crowds. This knowledge aids developers as it means they can focus their time and resources on the more important aspects, which influence how the virtual crowd is perceived. The related research has shown that using psychophysics has been successful in probing human perception, with regards to specific aspects of virtual crowd simulation. However, these studies have mainly focused on visual elements, such as graphics and animation. Additionally, as noted in Section 2.3, while the development and refinement of AI algorithms is prominent in research, evaluation, especially in the terms of perception or psychophysics, has limited saturation showing an avenue for further investigation.

In this thesis, a variation of these psychophysical methods are applied to specifically assess the behaviour of agents within an urban scene. With crowd behaviour being one of the most important factors for how a simulation is perceived (O'Connor et al. 2013*b*), psyhcophysical studies prove to be a useful method for enhancing the perception of crowd simulation. The psychophysical experiments conducted as part of this research are presented in Chapter 6.

3.3 Summary of Chapter 3

In this chapter, the history and methods of psychophysics have been outlined. It has been described how psychophysics began with Fechner and Weber through the JND and the Fechner-Weber law, only to then be refined and modernised over time. The various different methods from constant stimuli to staircase procedures have been detailed, noting their applications and key advantages in certain experimental situations. This is important due to the fact that the experiments this thesis presents in Sections 6.2, 6.3 and 6.4, utilise these psychophysical methods and will refer back to these highlighted methods and analyses. Additionally, related literature has been examined for perceptual evaluation and psychophysics, with reference towards crowd simulation where possible. This literature has provided evidence showing that perceptual evaluation, in particular psychophysical methods can be success at probing human perception with regards to facets of graphics and animation; however, it has also shown a gap towards more algorithmic elements in the form of crowd behaviour, an important part of crowd simulation that on many levels requires perceptual plausibility. This in essence shows what this research is contributing to the area. By applying psychophysics to quantify the subjectiveness of perception towards crowd behaviour, crowd behaviour can be objectively improved based on this perceptual feedback to become more plausible.

In the next chapter, the three-stage methodology of analysis, synthesis and perception is presented, noting the successes of the approach. This is the overall methodology applied in this thesis due to the robustness incorporated through the consideration of real-world data, which grounds perceptual evaluation in reality to provide meaningful results. Each stage is detailed, considering integral elements of the framework such as video analysis and the chosen psychophysical methods. In addition, by outlining the completed iterations of this methodology, a timeline of this research is provided.

Chapter 4

Methodology

In this chapter, the main methodology employed in this research is identified as the three-stage iterative method of analysis, synthesis and perception. The advantages and previous successes of this method are examined and linked back to provide a rationale. In addition, the intricacies of each stage are detailed, highlighting both the implications and processes involved in the context of this research.

4.1 Framework

For conducting perceptual experiments, it is useful to employ a specific methodology to ensure that it is kept grounded in reality (Peters et al. 2008). When probing human perception, the results are often highly subjective and require an element to keep them fixed within the realms being tested. To this end, a methodology consisting of three distinct stages has been employed for perceptual evaluation work towards virtual crowds (Peters & Ennis 2009). In this research, the plausibility of pedestrian group ratios were considered and perceptual evaluated. The first stage required the analysis of crowd data to inform the modelling process. The second stage used this data to synthesise crowd behaviours using a simulation model. The final stage then used stimuli from the simulation to perceptually evaluate the pedestrian group ratios. Results from this study were positive and applicable to informing the future creation of crowd simulations. This was in part due to this adherence to this methodology, which is noted to have some distinct advantages when applied to studies involving perception, see below:

1. By analysing real-world data the study is grounded in reality and unexpected results that might not be considered can potentially be highlighted.

2. This corpus of data can inform environmental and algorithmic construction to produce a tailored simulation for focused evaluation.

3. Through perceptual evaluation the impact of the synthesised crowds on viewers' judgements of realism can be established to inform future design.

These are important considerations for any perceptual study and given the successes present in the literature, an adapted version of this methodology was selected for this thesis. The three-stages cover, analysis, synthesis and perception. The tasks: (1) analyse real-world instances of crowd behaviour, (2) synthesis of crowd behaviour into simulation, (3) psychophysical evaluation of the resultant crowd behaviour. As noted in the literature, conducting these three stages is a necessity for producing a robust study that contains a model, which is ground in reality. By using this method over multiple iterations, it is possible to build a corpus of perceptual data that can be applied to various crowd behaviours.

Figure 4.1 shows the flow of this methodology beginning with the analysis stage, which is linked to the identification of a behavioural feature through crowd data. This information is passed onto the synthesis stage, which is also influenced by any perceptual data previously built-up within the corpus. The crowd simulation is refined and the behavioural feature implemented for the perception stage, where using psychophysical experimentation the behavioural feature for this iteration is evaluated. Due to the iterative nature of this adapted methodology, the perceptual data obtained is then added to the corpus for the next synthesis stage and the overall process cycles back to the analysis stage to start another iteration.

The methodology revolves around the acquisition of perceptual data through psychophysical experimentation, with regards to behavioural features. These behavioural features are implemented within a specially developed urban crowd simulation, which is discussed in detail in Chapter 5. Initially, this crowd simulation starts off with only the core algorithms as outlined in Section 2.1. This is due to the fact that the crowd simulation is being constructed with additional algorithms that refine the crowd behaviour over the iterations of this methodology. This is important, as it ensures the perceptual evaluation

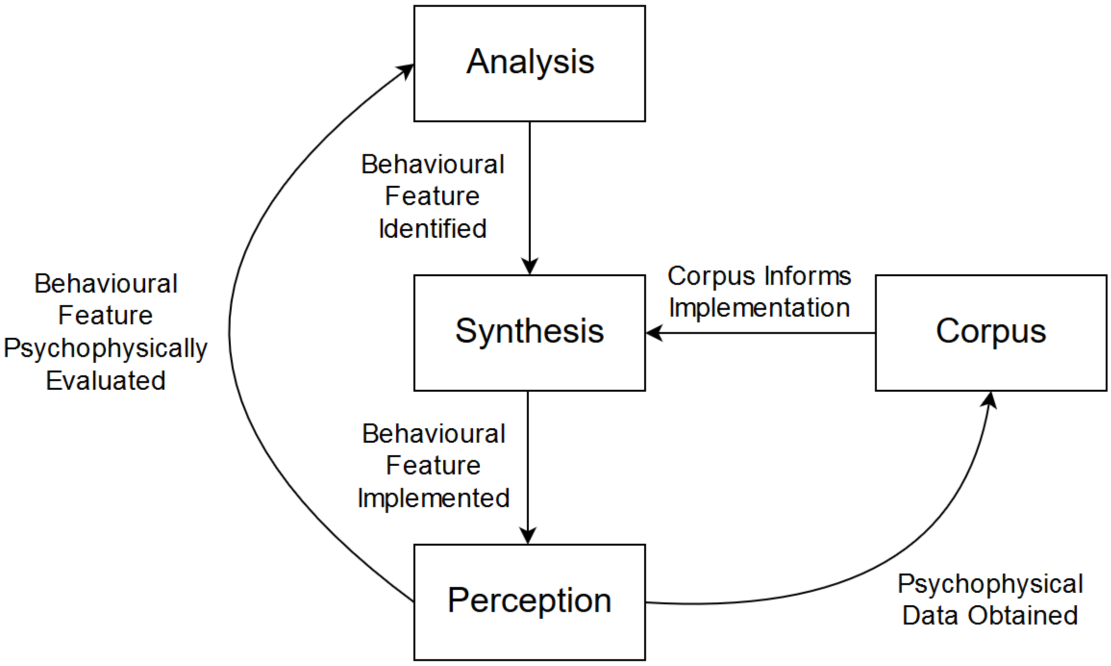

FIGURE 4.1: The adapted form of analysis, synthesis and perception employed in this research. The overall flow can be seen with the perceptual data being built-up into corpus over multiple iterations.

can be focused on a specific element without other areas potentially influencing results. In addition, it means that sophistication is added to the urban crowd simulation as a slow methodical process tempered by evaluation rather than by combining multiple AI algorithms, which as noted in Section 1.1 can fall prey of the 'complexity fallacy' and produce worse results.

Algorithm construction and modelling are informed by analysis of video data that shows real-world crowd behaviour within the location of interest. On occasion certain behaviours that are not predicted by the developers are discovered through these visual analyses. This can lead to a more robust simulation for perceptual evaluation. Modelling a simulation based on crowd data is still a fairly uncommon process, but there are a growing number of examples of this type of research shown in the literature examined in Sections 2.3 and 3.2. In addition to informing the modelling process for the simulation, the crowd data also highlights possible behavioural features that can be perceptually evaluated and their initial values, which can then be altered to produce the needed stimuli. The specifics of the analysis stage are discussed in Subsection 4.1.1.

Using the data obtained from analysis, it is then possible to model algorithm construction and the virtual environment. In a process of synthesis, the behaviours are created with the parameter spaces left open for perceptual evaluation to determine the optimum. Using values drawn from the real-world instances, the behaviour is initially based on reality before being subtly altered for psychophysical stimuli. Additionally, the virtual environment and behavioural annotation can be modelled based upon this data, taking into account the most important features that have an impact on crowd behaviour. More details on synthesis are discussed in Subsection 4.1.2.

By conducting a psychophysics experiment on the synthesised behaviour, it allows for the calculation of perceptual thresholds through psychometrics and identification of the most ideal intensities for implementing that behavioural feature. The specific psychophysical method employed varies based upon the behavioural feature. The compiled results from these experiments can then inform synthesis on future iterations of research and other related crowd simulations looking to have perceptually plausible behaviour. The perception stage is discussed in Subsection 4.1.3.

The purpose of applying this adapted methodology to crowd behaviour is for the accumulation of perceptual data. This includes optimum configurations for achieving perceptual plausibility for specific behavioural features, along with the threshold values for when a feature begins being perceptually realistic and when that is lost. This kind of knowledge is important for crowd simulation development as a whole, as it can inform which algorithms to implement and how to construct their parameter spaces to ensure perceived realism. In addition, by using threshold values it is possible to tweak both behaviour and algorithm design while keeping within the range of what is acceptable. This can be useful for both dynamic and outlier behaviour, as well as those simulations that require a tight leash on performance. Finally, due to algorithm construction being modelled on the reality presented in the crowd data, it is possible to analyse the differences between what the crowd data suggests is realistic and what the subjectiveness of human perception considers realistic. This highlights important links for consideration when perceptual plausibility is a key element in the simulation.

This thesis consists of three iterations of this methodology, encompassing three different behavioural features that are identified through crowd data, implemented into simulation and psychophysically evaluated. These behavioural features are varying velocity, social forces and grouping dynamics, respectively. The specifics of each with regards to the stages of analysis, synthesis and perception are presented in Chapter 6.

4.1.1 Analysis

Crowd behaviour in general may appear to be fluid in nature and even though some virtual crowds have been modelled using fluid dynamics, it is more common that crowd behaviour is modelled using multi-agent systems (Kapadia et al. 2012) because it allows for control over individuals and in turn flexibility to implement the unexpected behaviours that are present in reality. When developing behaviour, it is possible for some behaviours to be overlooked, e.g. emergent behaviours such as abrupt changes in direction or spontaneous ceasing of motion. In the analysis stage, real world crowd behaviour is considered in order to highlight these potentially unexpected behaviours and inform the development of the virtual crowds and environments. Over time and with multiple iterations of the methodology, a corpus of analysed perceptual data can be built-up. In this thesis the main medium considered for analysing crowd behaviour is video footage, showing actual pedestrians in a variety of urban locations. It is possible to use other mediums such as photographs, however video gives a more accurate representation of the motion, grouping and behaviour of the pedestrians. This can lead to a more robust simulation for improved stimuli generation, which is then utilised for psychophysical evaluation.

4.1.1.1 Behavioural Features

The identification of a behavioural feature is an important aspect for this research and a key element of the analysis stage. A behavioural feature provides a focal point for the analysis and the eventual perceptual evaluation. A behavioural feature is an aspect of the virtual crowd, which influences its crowd behaviour in some manner. Over multiple iterations of this methodology, more and more behavioural features are identified offering key ideas about how to implement virtual crowds. The most apparent behavioural feature that has not already been examined and is shown prominently within the crowd footage is selected. In this manner the most common features are evaluated first, as it can be argued they have the most initial value when producing crowd behaviour. Over time however, more specific and less prominent behavioural features would also be analysed. In addition, it is possible for a single behavioural feature encompass multiple elements, for example social forces would consists of multiple forces as part of one behavioural feature. This thesis presents three behavioural features for analysis, synthesis and perception, as outlined in Chapter 6.

4.1.1.2 Video Analysis

The crowd footage was acquired at peak times from live web camera feeds that were a part of the EarthCam network. Each of the locations featured high numbers of pedestrians, which allowed for the collection of useful of behavioural data to later influence simulation development. These urban locations consisted mainly of high pedestrian trafficked areas such as pedestrianised streets, which are a common type of urban location modelled for crowd simulation. They feature prominently in virtual worlds presented in both video games and serious applications. This means that by analysing this type of crowd and location, the resultant behavioural data would be of most use for future implementations and applicable to a wide range of related research.

With the identification of a behavioural feature for analysis, it means that only this specific behaviour now needs to be taken into account. Previously, viewing the overall footage has highlighted different and unexpected behaviours that can be used to inform simulation design, but for the final part of the analysis stage a manual annotation is tailored towards the behavioural feature. Through visual identification at key frame intervals, annotations are made with their relevant variables recorded. Figure 4.2 shows some examples of the camera feeds available on the network. The specifics of each video analysis are detailed as part the relevant experiment Sections in Chapter 6.

4.1.2 Synthesis

The synthesis stage sees the utilisation of the data gathered during analysis to inform development of a tailored crowd simulation for perceptual analysis. The identified behavioural feature is implemented within this urban crowd simulation; however multiple configurations are required for the purposes of psychophysics. This means the normal modelling approaches have to be altered to allow for parameter spaces and customisability. Each behavioural feature variable must be implemented in a manner to replicate the reality present in the crowd data, while being alterable to allow for other configurations

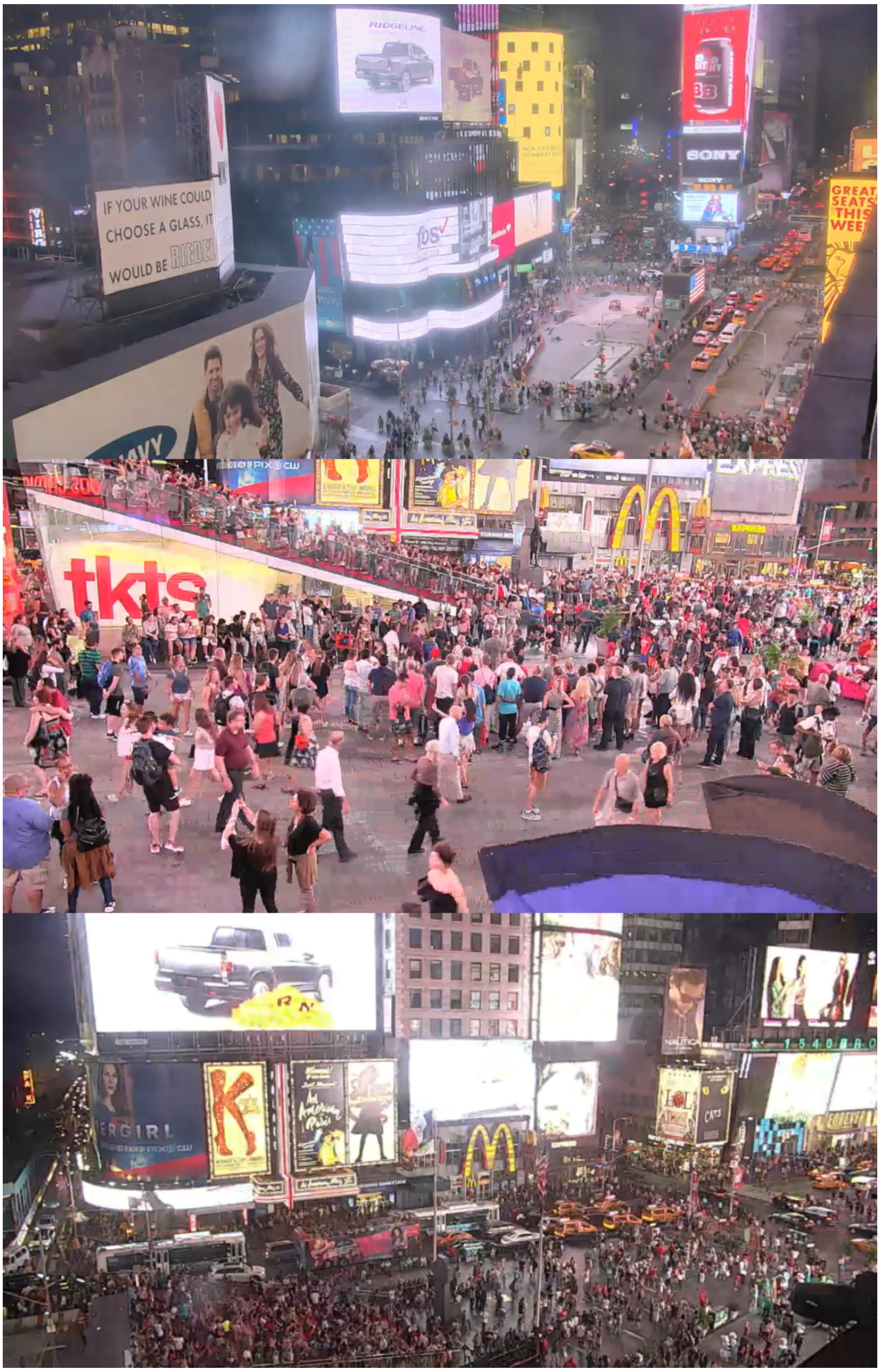

FIGURE 4.2: An example of the live camera feeds around Times Square.

to generate stimuli for psychophysical experimentation. In this stage, the general refinement of the urban crowd simulation platform takes place, with upgrades to graphics, lighting and general performance being common. The output from this stage is a series of generated videos showcasing the behavioural feature implemented for that iteration.

The videos each contain different intensities of the stimuli and are used in the perception stage. Further information regarding the implementation of the urban crowd simulation and by extension the process of synthesis is outlined in Chapter 5.

4.1.3 Perception

The perception stage of the methodology is primarily to perceptually evaluate the identified and then implemented behavioural feature. This provides feedback regarding the perceived realism of the feature and can identify key intensities or parameter values that appear most realistic to viewers. The literature examined in Section 3.2 has shown that using perceptual methods is a viable approach for highlighting what viewers find perceptually plausible. It is possible to test a number of different variables or a whole simulated scene, however while this can provide a useful indication of how a combination of factors can affect the perceptual realism, it does not necessarily judge the merit of the individual components. This is where testing a specific behavioural feature is important, as it not only highlights the importance of that specific factor in comparison to others for modelling purposes, but perhaps the best configurations when implementing said feature in a number of scenarios.

In order to achieve this, psychophysical methods are applied. Stimuli are selected and altered in intensity and the response from viewers is measured to determine certain perceptual thresholds through psychometrics. Using this type of a study allows for the behavioural feature to be effectively assessed with regards to its perceived realism. Participants are presented with a number of short video clips, showing variation with regards to the behavioural feature. This variation is in the intensity and varies depending upon the specific composition of the behavioural feature. The order and configuration of the video clips is determined by the individual experiment and the psychophysical methods being employed. The final results from the experimentation are added to the corpus of perceptual data and a new iteration of the methodology begins. The three psychophysical experiment conducted in this research are presented in Chapter 6.

4.2 Summary of Chapter 4

In this chapter, the iterative methodology of analysis, synthesis and perception has been discussed. The advantages of employing such a methodology for perceptual evaluation have been highlighted. These advantages include being able to ground the subjective nature of the perceptual data reality to ensure its applicability. In addition, over the course of multiple iterations a corpus of perceptual data for different behavioural features can be built-up and synthesised to inform future experimentation. The intricacies of each stage have been examined, whereby in analysis video data is analysed for behavioural feature identification and in synthesis it is implemented into simulation. Finally, in perception the behavioural feature is employed as stimuli for psychophysical experimentation. By adhering to this methodology, three behavioural features have been identified, implemented into the urban crowd simulation and evaluated through psychophysics.

In the next chapter, the implementation of the urban crowd simulation is outlined. The purpose for bespoke development is considered, with regards to perceptual evaluation and psychophysics. An overall look at the development cycle is provided to give an idea of the gradual flow of work in accordance with the synthesis stage of the general methodology. The implementation of the different elements that compose the urban crowd simulation are covered in-depth. The core algorithms, behavioural annotation, social forces, Unity platform and grouping dynamics are considered and linked back to the literature to provide justification for the specific development choices.

Chapter 5

Implementation

In this chapter, the purpose of developing a bespoke urban crowd simulation is explained. Low-level details regarding the development cycle of the urban crowd simulation are provided, from the initial application using OpenGL to the final iteration in the Unity environment. Special note and attention is given to the implementation of the key behavioural features, namely varying velocity, social forces and grouping dynamics.

5.1 Urban Crowd Simulation

Virtual crowds can be implemented using a number of different methods and algorithms, as discussed in Chapter 2. Often the purpose of employing a crowd simulation is to replicate real-world behaviour in a given scenario to provide forms of interaction and in some instances useful data (Almeida et al. 2013). This research is examining crowd simulation in general, through the use of psychophysics to evaluate crowd behaviour. As such, the type of crowd simulation required does not fall under typical purposes. An average crowd simulation simply has a configuration selected by the developer and that is that. Psychophysics however, requires the potential for different configurations and not only the ones that appear most suitable to the developer. The main purpose for developing the urban crowd simulation utilised in this thesis was to allow for this needed flexibility.

Different intensities of stimuli are required for psychophysical evaluation. By developing a bespoke urban crowd simulation, parameter spaces and algorithmic alteration could be made to allow for customisability. By incorporating variables and weight factors to be passed through algorithms to control the level of influence over crowd behaviour, it is possible to alter values to create different intensities for a specific behavioural feature during runtime. As such, by running the simulation through multiple iterations and each time altering these values, it was possible to generate the different intensities of stimuli for experimentation. This shows the need for specially developing the urban crowd simulation, as it required an element overlooked for typical implementations.

The iterative methodology was outlined in Chapter 4 and gives another purpose for developing the urban crowd simulation specifically for this research. The synthesis stage is the point at which newly identified behavioural features are implemented into simulation. The adding of sophistication to the crowd behaviour is a slow and methodical process, which is influenced through the perceptual data accumulated in the corpus. The initial development of the urban crowd simulation began with the core algorithms as highlighted in the literature, see Section 2.3. Using this as a base for the implementation of behavioural features, the initial crowd behaviour would be devoid of other influencing factors. This is an important consideration for the accuracy of the results from the psychophysical experiments. By gradually adding behavioural features over time and tempered with the perceptual data from previous experiments, the stimuli can be better focused towards the specific behavioural feature required for study.

5.2 Development Cycle

The development cycle of the urban crowd simulation was an iterative one, through the application of the synthesis stage of the general methodology. Sections 5.3 to 5.9 look at specific elements of the development cycle in detail, however an overview of flow of implementation and how it fits with the methodology is important.

The entire development cycle was conducted over three iterations of synthesis. Initially, the core algorithms were implemented to form the basis for the virtual crowds and then at each iteration a new behavioural feature was implemented into the urban crowd simulation. In addition, other refinements took place over the course of synthesis. The initial development environment utilised the C++ programming language for algorithm design, in unison with OpenGL for the graphical representation; however, during the third iteration the development environment was changed to the Unity game engine.

The core algorithms were implemented in the first iteration of synthesis, along with the varying velocity behavioural feature. The second iteration of synthesis saw the addition of the social forces behavioural feature. Both the first and second iterations of synthesis took place within the C++ OpenGL development environment. During the third iteration of synthesis the switch to the Unity development environment took place and the grouping dynamics behavioural feature was implemented. Figure 5.1 gives an overview of the development flow with respect to the main AI algorithms. The virtual environment that was implemented for C++ OpenGL used a procedural method to generate a cityscape with functionality for behavioural annotation, whereas recreations of actual locations were implemented for Unity with its own version of annotation. These differences in virtual environments are due to the specific methods employed for the different experiments.

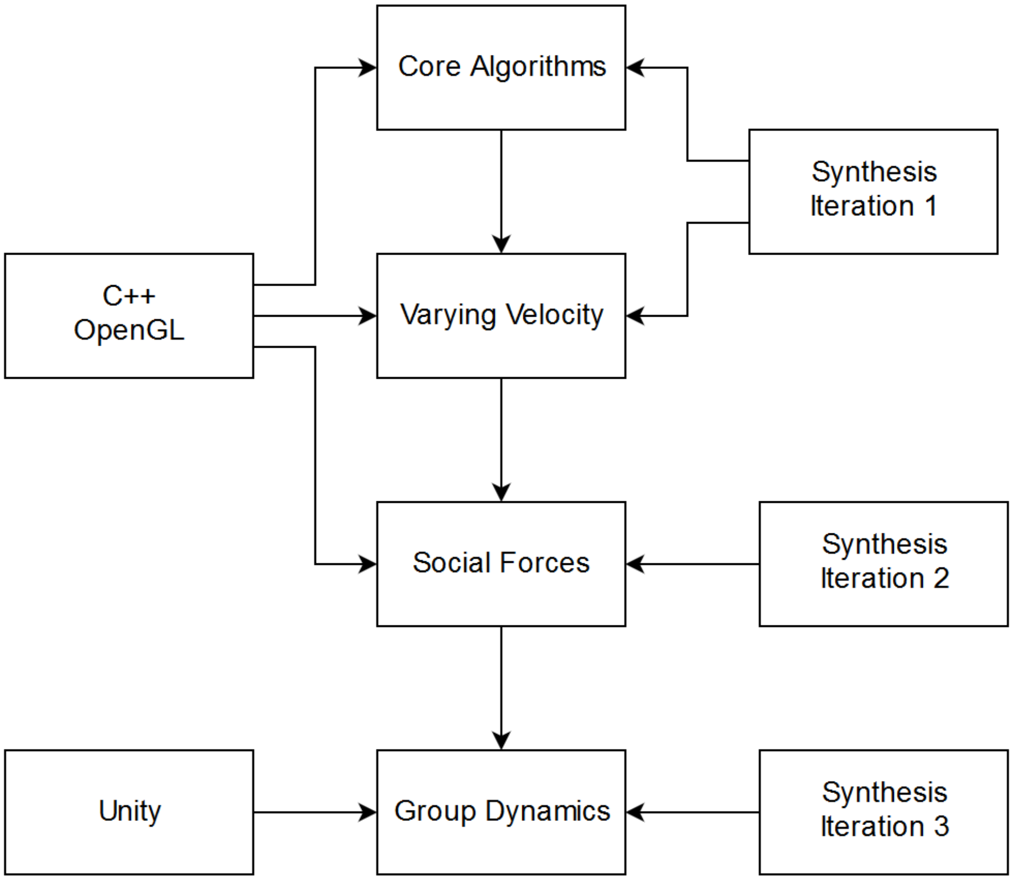

FIGURE 5.1: An overview of the development cycle for the main algorithms, noting the development environment on the left and the general methodology iteration on the right.

The initial development environment using C++ and OpenGL was selected due a number of considerations. C++ is a middle-level programming language that has support for a large number of platforms and allows for object-orientated programming, which was of benefit for virtual crowd implementation. In addition, C++ was compatible with OpenGL and the OpenGL Utility Toolkit (GLUT) for graphics programming and window visualisation. The main advantage of using this combined development environment however, was the flexibility it presented for algorithm construction and thus allowed for customisability in terms of the behavioural features.

The move to Unity was due in part to the general methodology. As outlined in Section 4.1, the corpus of perceptual data gained through previous iterations of the methodology influence the synthesis in the current iteration. It was suggested within the perceptual data, that the basic geometric representation of the agents and their animations could be obscuring the small changes in speed and direction. By changing the development environment to Unity, the agents could be updated with higher poly-count models and realistic animations. Games technology has been utilised in the past for simulating crowds (Szymanezyk et al. 2012) and in addition to the graphical improvement, Unity allowed for the creation of UI elements to implement behaviour and other elements in a more efficient and straightforward manner.

5.3 Core Algorithms

For psychophysical evaluation, it was important to begin with a base from which behavioural features could then be incorporated to tailor crowd behaviour. By analysis of the different AI algorithms used for crowd simulation in Sections 2.1 and 2.2, in addition to the examination of relevant literature in Section 2.3, it was possible to see a trend of core algorithms that would allow for just that. By implementing the framework of a decision making system, pathfinding navigation and local steering mechanics, it would allow for the basic operations of an agent, namely to perceive, think and act (Anderson 2003). For an individual agent, the decision making system selects the destination, pathfinding calculates the path needed to reach that destination and steering makes the local adjustments and collision avoidance using synthetic perception to get it there successfully. Figure 5.2 shows this process of simulating crowd behaviour with the core algorithms.

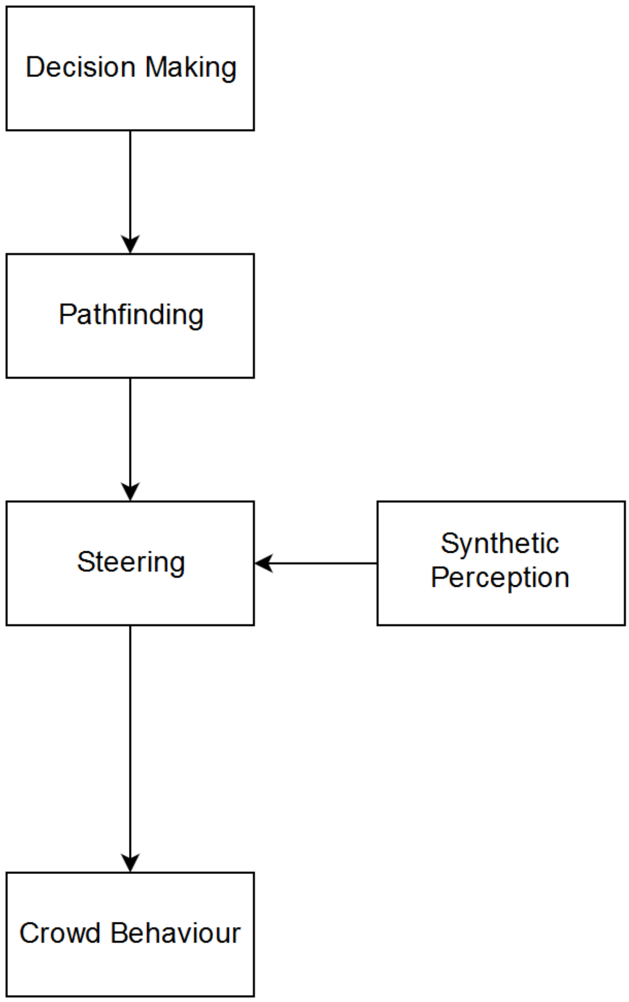

FIGURE 5.2: A diagram showing the communication between the core algorithms, with the resultant output being crowd behaviour.

The different types of crowd simulation are outlined in Section 2.1. The agent-based approach was selected for the advantages it possessed over both flow-based and entity-based crowd simulations. These advantages are derived from the fact that each agent is treated as an individual, allowing for more complex behaviours beyond the fluid and physics based models of the other crowd simulation approaches. The literature examined in Section 2.3, also reveals that multi-agent systems are currently prominent for both research and virtual crowd implementation. Agents within the urban crowd simulation are treated as individuals, each with their own variables and data structures influenced separately by the AI algorithms to collectively produce dynamic crowd behaviour. Each agent is updated for each frame of the simulation runtime.

Finite-state machines were selected for the decision making system, due to the fact only a limited number of states were required for implementation. This makes the algorithm effective in terms of performance and due to the nature of extensibility the algorithm possesses, it is ideal for extension in future to add more states for identified behavioural features. An initial state was implemented for path following, whereby the decisions are made for selecting between different destination nodes based upon visible annotations within the virtual environment and a random-chance factor. In this way the overall behaviour is kept fairly simplistic, which is ideal for psychophysical evaluating a specific behavioural feature and agents still function in an expected manner.

The A* algorithm was selected for the pathfinding element as it has proven robust (Cui & Shi 2011) and has some advantages that are of benefit for crowd simulation. It guarantees to find the shortest path by searching in the direction of destination node, but can backtrack if that is unsuccessful to find an alternative route. Additionally, with a node system set-up effectively its drawback in terms of efficiency if the path is not easily locatable is eliminated. Due to the virtual locations in this research being pedestrianised urban areas, this does not pose an issue. For the urban crowd simulations implementation, an underlying grid was installed within the virtual environment. This grid consisted of nodes and their connections, with each connection containing data as to whether a path between two specific nodes is walkable or not. The output from the A* algorithm is a list of viable nodes that an agent can iterate through in the process of movement, leading to the destination node as selected by the decision-making algorithm. The algorithm for this implementation of A* can be seen in Equation 5.1.

$$f = g + h \tag{5.1}$$

Where f is the calculated fitness value for a specific node. This fitness value is an estimate for the cost of a path going through that node. g is the goal value, which is the cost from starting node to this node. h is heuristic value, which is the estimated cost from this node to the destination node, using the Manhattan method. A* begins at the starting node and calculates a fitness value for the surrounding nodes. It then moves to the node with the lowest fitness value and repeats until the destination node is found. The path is then traced back through the nodes with the lowest fitness values, providing the shortest path to the destination.

In terms of steering, the crowd path following model was implemented for the urban crowd simulation (Reynolds 1999). In addition, a radial approach for synthetic perception was incorporated for the purpose of agents being able to access their local neighbourhoods, with respect to steering force calculation as seen in Figure 5.3. This specific steering mechanic was selected as it is the most relevant for crowd simulation. Boids (Reynolds 1987) for example is to simulate swarming, whereas crowd path following is a combined mechanic that achieves what it suggests, crowds following paths. It consists of two separate steering mechanics, path following and separation. The paths are passed to agents from the pathfinding algorithm and the path following steering force is calculated towards the current node. For managing the crowds and avoiding 'clumping', a separation steering force is calculated when agents are detectable by the synthetic perception as being within close proximity.

FIGURE 5.3: An example of radial perception and a seperation steering force (Reynolds 1999).

These core algorithms communicate starting with decision making, which passes a destination node to the pathfinding algorithm. This in turn outputs a path of nodes to said destination, which is then used by the steering mechanics to provide a steering forces towards the various nodes. Using radial perception agents avoid colliding with each other through the separation steering force. The implementation of these core algorithms provided a base of basic crowd behaviour for future extension with regards to behavioural features. This process of starting with a basic platform and then adding sophistication through analysing crowd data is important for the general methodology and the psychophysical experiments.

5.4 Behavioural Annotation

Behavioural annotation is a method rather than a specific algorithm, in which features within the environment are 'labelled' in such a manner that they can be detected by agents and cause a form of behavioural change. A simple example would be to label a road, which in turn causes agents to detect it and slow down when approaching it. A label can be something as simple as a Boolean value just to note the presence of a specific feature or it could contain more in-depth control information (Peters et al. 2003), such as a weight or an intensity. The important element of annotation however, is that this data is kept within the virtual environment rather as part of the agent systems. This has advantages in terms of performance and what can be accomplished with respect to behaviour. Behavioural annotation also makes a positive appearance in the literature examined in Section 2.3 and thus was an important design decision for future behavioural features.

A behavioural annotation system was implemented for both the C++ OpenGL and Unity Development Environment. The implementation varies, with C++ OpenGL using the underlying grid to store data with regards to annotations. For Unity the annotations such as vector fields and specific zones, are placed directly within the virtual environment itself as objects. Figure 5.4 shows an early prototype of the urban crowd simulation. The underlying grid is visible showing the annotations, with blue representing an annotation for the road and green representing an annotation for the pedestrian crossing. When agents detected these annotations their behaviour altered, with agents slowing down to cross roads and forming lines for crossing the pedestrian crossings. These annotations are not present past this prototype, as they were merely a test for the behavioural annotation system.

While the annotation systems were put into place, they were not specifically used for the identified behavioural features. This is due to the fact that the process of analysis identifies the most obvious features and behavioural annotation is better for tailoring specific behaviours towards features in the environment. As such, a compatible behaviour was not identified for the first three iterations, however with the system in place it can provide a focal point for future work.

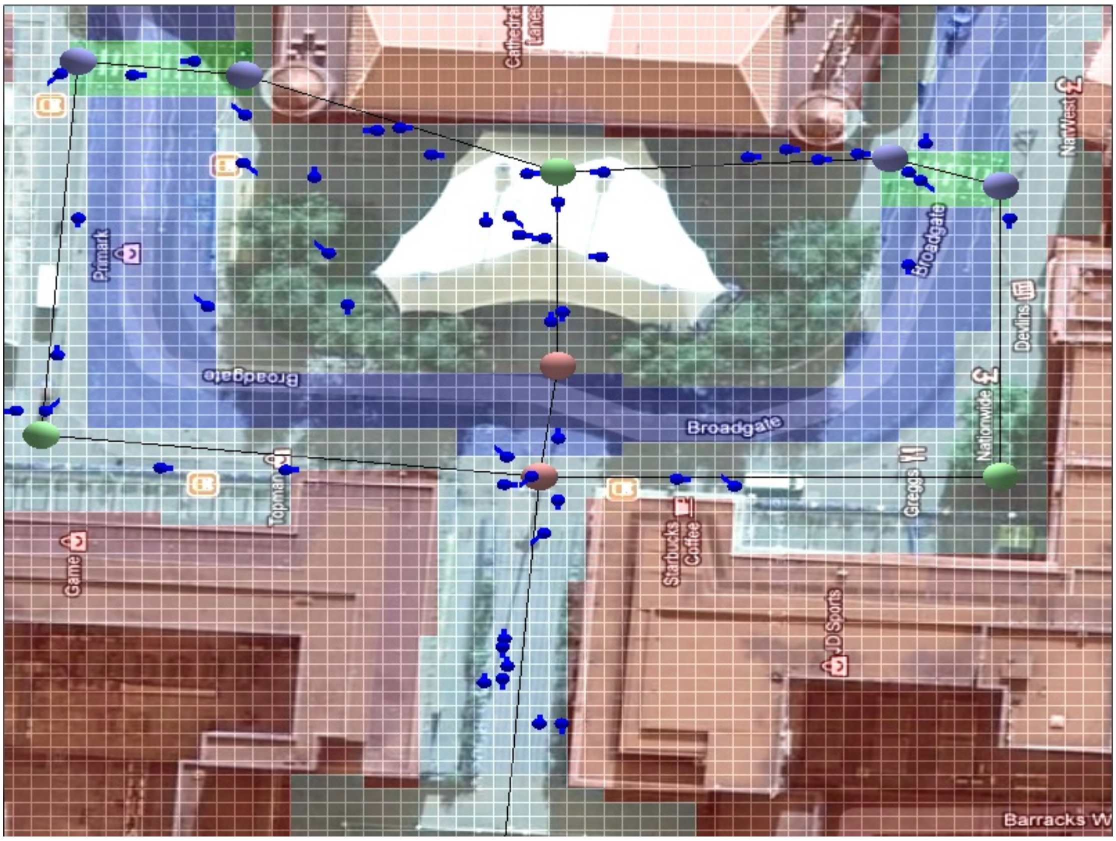

FIGURE 5.4: An early prototype of the urban crowd simulation that uses a 2D plane textured with a satellite view of Coventry City Centre for the virtual environment. Behavioural annotations are displayed within the grid, each colour representing a different type. Agents are shown as blue spheres, with a line representing their current driving force.

5.5 Procedural Environment

As noted in the previous Subsection, early prototypes for the urban crowd simulation used a 2D plane with a static texture for the virtual environment to test the core algorithms. Past testing however, this type of environment would not be suitable for the psychophysical experiments. A full 3D virtual environment was implemented through a procedural approach to generate a cityscape. Using this method allowed for the possibility of multiple configurations as required. The urban environment was automatically generated at runtime based on a series of predefined rules for general shape, structure and layout. It was possible to generate different layouts due to the parameterised nature of the model. Once a suitable configuration was found it could be generated through the specific seed. This is important as the environment needed to remain constant for stimuli generation over multiple runtimes. The main benefit of employing a procedural method for generating the virtual environment was that it allowed for substantial complex geometry in terms of the buildings and roads, without the requirement of programming all the graphical specifics via OpenGL.

The geometry that forms the roads and architecture for the virtual city is rendered through OpenGL and visualised with GLUT. The implementation is derived from an open source toolkit written in C++ (Lance et al. 2013). The procedural rules for the geometry elements and overall configuration are defined in the toolkit. Land generation rules are used for object placement and layout. These rules create templates via subdividing quads and triangles, which are then populated with the geometry that combined forms the urban architecture, which in turn forms the cityscape as generation progresses. Other elements are initialised as part of the procedural generation routine such as materials, light sources and camera controls. The resultant virtual city is approximately 100km and includes three zone types, commercial zones, residential zones and industrial zones. This gives a virtual urban environment, which can then be populated with agents running the core algorithms. Figure 5.5 shows the final generated virtual urban environment.

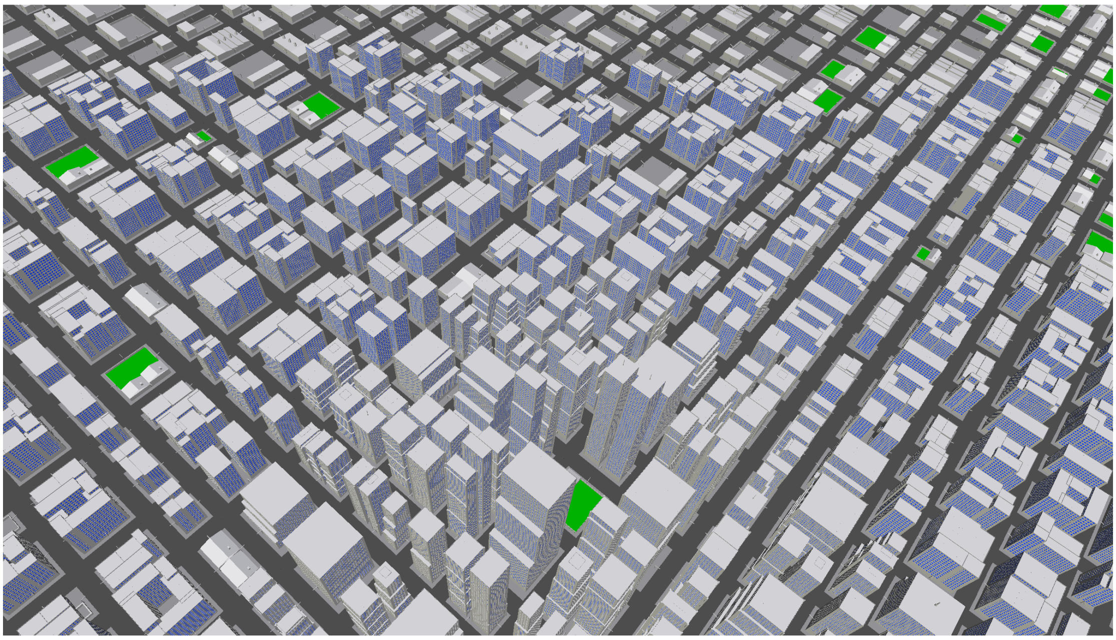

FIGURE 5.5: An example of the procedurally generated cityscape, showing the commercial, residential and industrial zones. The generated city while covering a large area, maintains an urbanised configuration through the defined procedural rules.

5.6 Varying Velocity

Velocity is an important consideration for behaviour and an element that is often overlooked. When considering crowds in general however, it is one of the first elements that is noticeable. Velocity consists of two components, a directional vector and a magnitude. In implementation terms, this is the driving force for the agent's movement, providing both a direction and a speed. An agent's final velocity therefore can be seen as the resultant directional vector and rate of motion calculated from the steering forces and any other movement altering factors present and applicable within the specific simulation. Figure 5.6 shows an example of agents with their driving force visible within the procedural environment. In this instance the varying velocity feature is active.

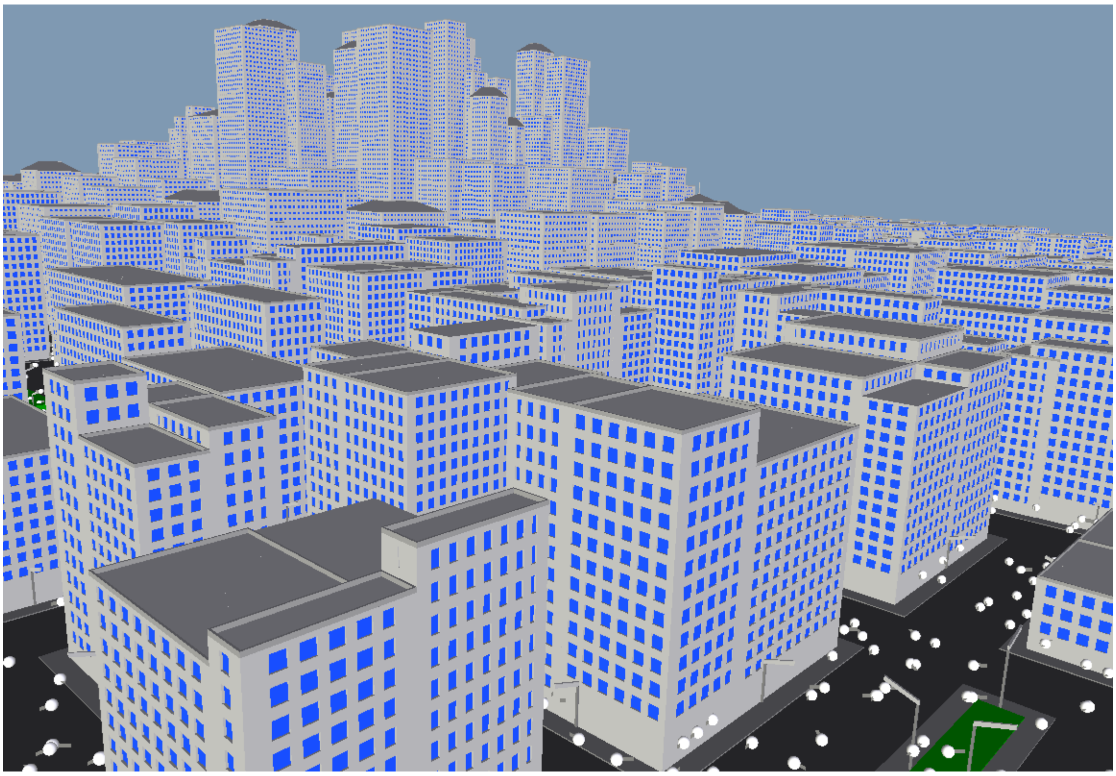

FIGURE 5.6: An example of the procedurally generated virtual environment populated with agents. The agents can be seen at street level as spheres, with their driving force visible as a line. Here they are under the influence of the varying velocity behavioural feature and as such are kept within the defined constraints.

For the urban crowd simulation, the varying velocity behavioural feature consists of two parameters used to restrict agent velocity, more specifically the magnitude or speed component, within certain limits. As velocity is the typical manner in which agents are given motion in most all simulations, the name was selected to show that this feature is applied towards that element. This behavioural feature fits into the algorithm design after steering mechanics have been applied and final velocity for the agents is calculated, as seen in Figure 5.7. The two parameters can be varied to allow for stimuli generation with respect to psychophysics. The two varying velocity parameters are:

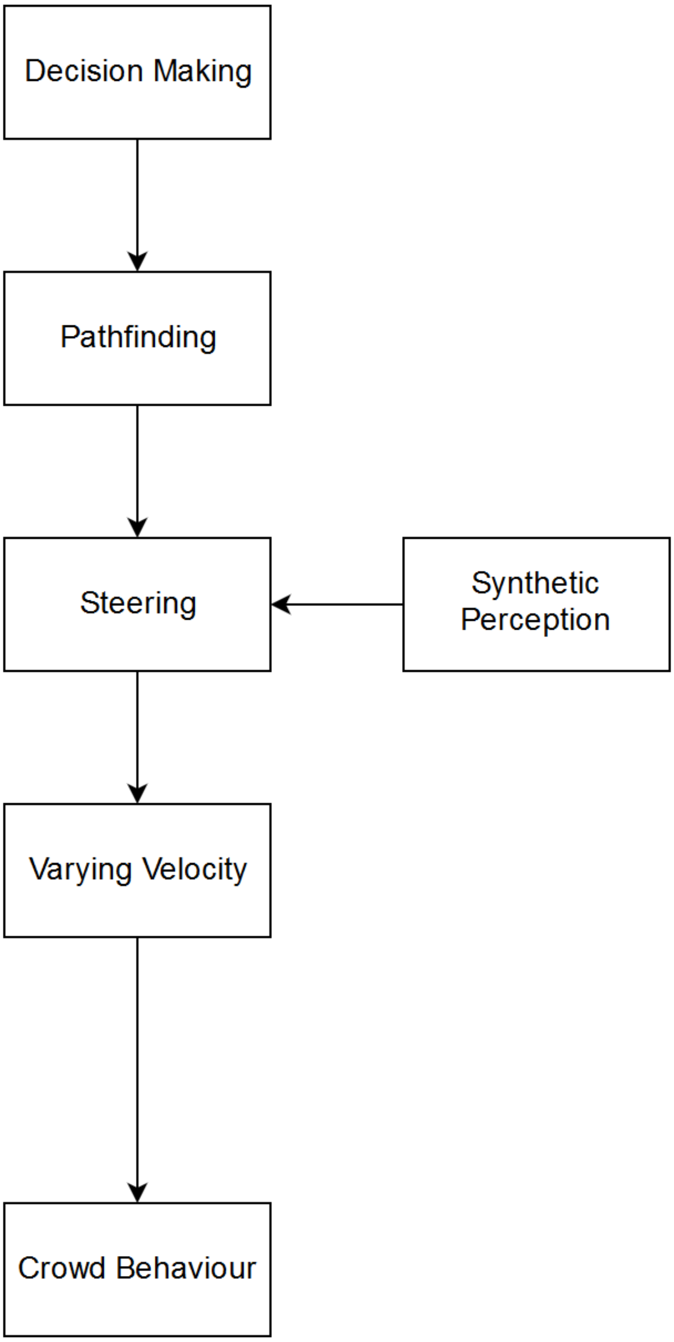

FIGURE 5.7: A diagram showing the communication between the core algorithms and the varying velocity behavioural feature.

1. **Velocity Range**: This parameter determines the minimum and maximum velocity magnitudes available to the agents, in essence the different speeds they can

operate within. For example, agents can be restricted to a range of speeds between 0.5 m/s and 2 m/s for the purposes of their behaviour.

2. **Velocity Distribution**: This parameter determines the distribution for the average magnitude utilised within the velocity range. When referring to agents, this is often called the preferred velocity and is essentially the speed that agents will try to maintain.

With these two parameters making up the varying velocity behavioural feature, it is possible to give agents a velocity range or speed limits to keep within, while also giving them a velocity distribution or a preferred velocity, so that they try to maintain a certain speed within that range. As such by altering these parameters, it was possible to generate different intensities of stimuli. By implementing a scale between 0 and 1 for these parameters based on the average velocity magnitude or walking speed identified via real-world crowd data, the intensity could be controlled in clear steps or percentages, which is important for psychophysical evaluation. The specifics of the psychophysics experiment for the varying velocity behavioural feature are presented in Section 6.2.

5.7 Social Forces

Social forces are an important element when considering crowd behaviour. Multiple implementations were noted in the crowd simulation literature covered in Section 2.3, however potentially the most fundamental and the one utilised as the basis for this research is Helbing's original social forces model (Helbing & Molnar 1995). Social forces can be used to describe the internal motivations or 'social forces' of the different agents. These social forces can be seen in real-world crowds, thus showing the importance of such a model for virtual crowd behaviour. A typical implementation consists of several different forces (Mehran et al. 2009), which are calculated to produce a steering force that is applied to an agent's driving force. Helbing's social forces model is the same, employing a force towards a desired velocity, coupled with attractive and repulsive forces from other agents and objects. The final result is often convincing crowd behaviour, with the different forces leading to a style of self-organisation that appears organic in nature.

For the urban crowd simulation, a social forces behavioural feature was implemented based on Helbing's model using three key forces:

1. **Agent Repulsion**: The force of repulsion between agents. This force is calculated between agents to produce a steering force that causes certain agents to move away from other agents.

2. **Agent Attraction**: The force of attraction between agents. This force is calculated between agents to produce a steering force that causes certain agents to move towards other agents.

3. **Object Repulsion**: The force of repulsion between an agent and an object. This force is calculated between agents and specific objects to produce a steering force that causes certain agents to move away from specific objects.

These three forces were selected as they form part of the fundamental basis for social forces and the Helbing model, however object attraction is an additional force present in the Helbing model but not considered in this implementation. This is due to the model being stripped back to control emergent behaviour for the relevant experiment, which aimed to explore the forces between agents. The object repulsion force was included as a means to control collision detection between agents and specific scene geometry.

Each force was implemented with a weighting value as part of its parameter space, meaning that the intensity of each force could be varied for stimuli generation and then utilised in psychophysical evaluation. Equation 5.2 shows the algorithm implementation and how the overall steering force is calculated from the three individual forces.

$$F(t) = w_{aa}F_{aa}(t) + w_{ar}F_{ar}(t) + w_{or}F_{or}(t) \tag{5.2}$$

Where $F(t)$ is the calculated social force at that time step, which is used as a steering force for the agents. The agent attraction force at that time step is represented as $F_{aa}(t)$, the agent repulsion force at that time step as $F_{ar}(t)$ and the object repulsion force at that time step as $F_{or}(t)$. The individual force weights are represented as the modifiers w_{aa}, w_{ar} and w_{or}. These different forces calculated indivdually can be seen in Equations 5.3, 5.4 and 5.5.

$$F_{aa}(t) = rN\sum_{i=0}^{n}(a_t - a_{it}, t) \tag{5.3}$$

$$F_{ar}(t) = rN \sum_{i=0}^{n} (a_{it} - a_t) \tag{5.4}$$

$$F_{or}(t) = rN \sum_{i=0}^{n} (o_{it} - a_t) \tag{5.5}$$

Where a_{it} is the position of all the agents within the current agent's synthetic perception and o_{it} is the position of all objects within the current agent's synthetic perception, at that time step. The current agent's position at that time step is represented as a_t. A normalising factor is represented as N and r represents the behavioural fluctuation factor, which adds some of the minor variation possible in behaviour. t represents the current time step and in the case of the agent attraction force, this shows that attractiveness of the agents currently being considered starts high but typically decreases over time, as the agent's interest declines. For the repulsive forces, however, the time step in and of itself doesn't influence the force calculation. Note, in terms of the synthetic perception, agents in the direction of the desired location are perceived more highly for force calculation.

The output of the social forces algorithm is a steering force, which is then communicated to the steering mechanics algorithm for consideration in an agents final calculated velocity or driving force. Figure 5.8 pictures an example of agents within the procedural environment with social forces active. Figure 5.9 shows the communication between social forces and the core algorithms.

The parameter weight value is what allows for different intensities with regards to the individual forces. The weight value is passed as a float between 0 and 1, allowing for a percentage influence of each force with the initial average values based on crowd data. The psychophysics experiment on the social forces behavioural feature is presented in Section 6.3.

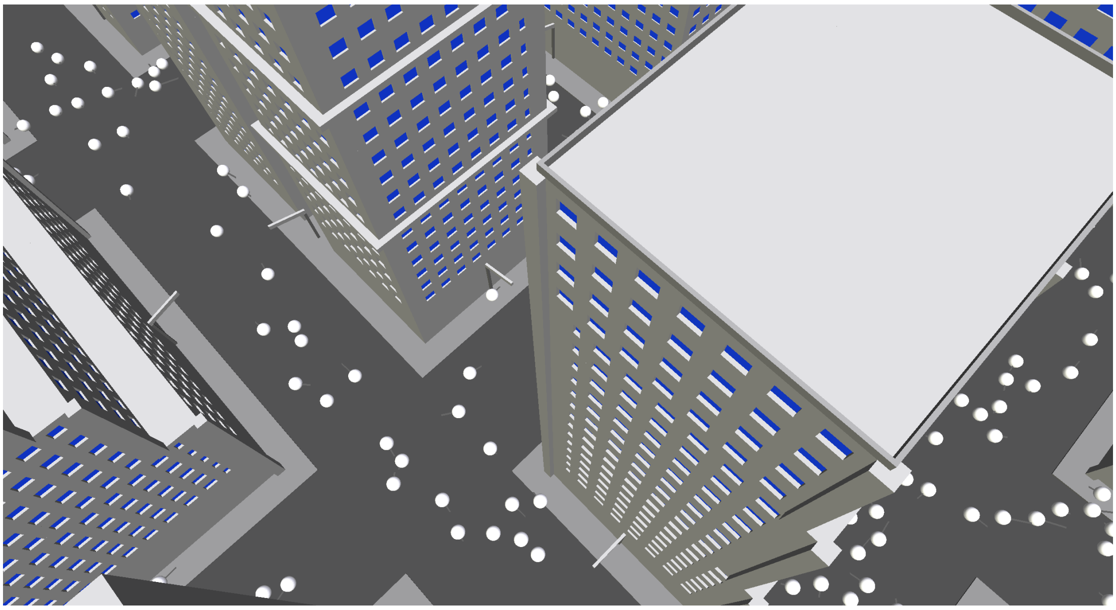

FIGURE 5.8: An example of agents within the procedural environment, with the social forces behavioural feature in effect. In this instance, the effects of the three social forces can be seen in the behaviour of the agents and the overall positioning reflects self-organisation.

5.8 Recreated Environments

The urban crowd simulation was changed to the Unity development environment to utilise the benefits as outlined in Section 5.2. The third iteration's identified behavioural feature was grouping dynamics, as discussed in Section 5.9. Three urban locations were selected from a variety of the crowd footage. For the purposes of psychophysical evaluation, the virtual environments needed to closely resemble the real-world locations. Unity utilises a system of prefabs, whereby game objects with their defined properties can stored in this manner, to act as a template for future use within a scene. In this way multiple instances of the same object can be added to the scene, inheriting their initial properties from there prefab. By producing prefabs for common urban objects and architecture, the three scenes required could be efficiently built from these templates. The main consideration was ensuring that the areas available for navigation by pedestrians closely matched those presented in the footage. By using stills from the footage as a visual guide, the virtual scenes were built from the ground up using these prefabs. Figure 5.10 shows the construction of Bourbon Street virtual environment within Unity. Figures 5.11, 5.12 and 5.13 show the created virtual environments with their real-life counterparts for comparison.

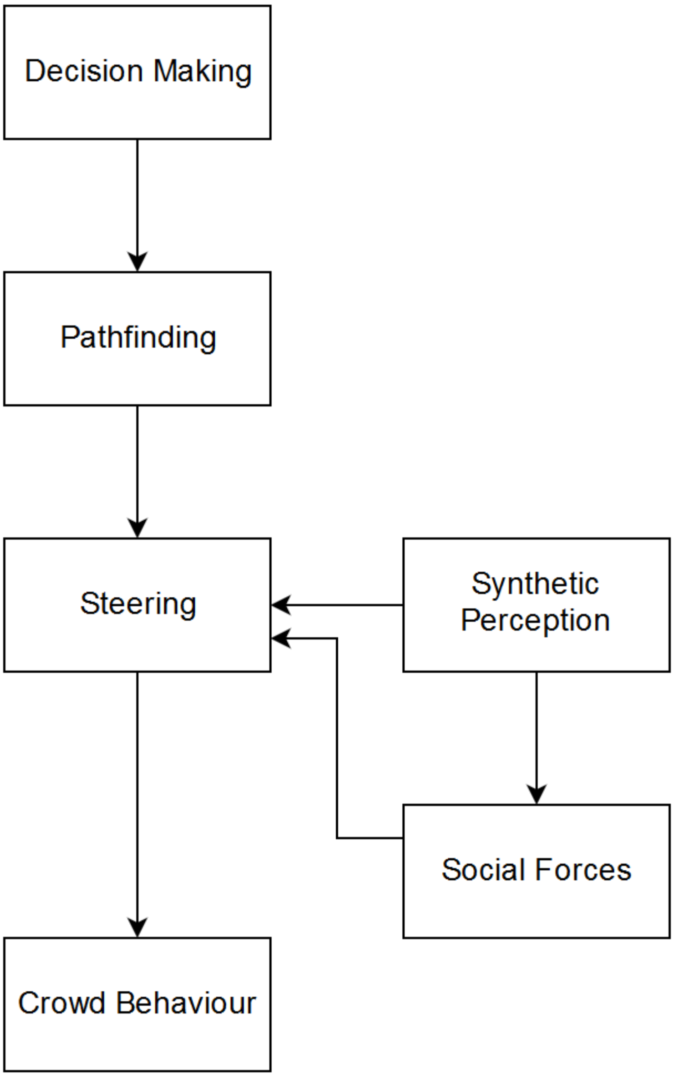

FIGURE 5.9: A diagram showing the communication between the core algorithms and the social forces behavioural feature.

The virtual environments provide the 3D geometry in which the agents could be generated, however other properties had to first be implemented into the environment to ensure that the agent systems behaved appropriately. The walkable areas and inaccessible areas needed to be communicated to the agents so that they kept to the locations required and did not collide with elements of the geometry. Through a form of annotation, vector fields of various magnitudes were placed into the virtual environment. This caused agents to adhere to the walkable areas, while still allowing them the freedom of movement required by their steering mechanics. These vector fields remained the same over all runs of the urban crowd simulation for stimuli generation.

The paths that agents followed were defined within the virtual environment. Using a series of nodes and their connections, as is common for pathfinding algorithms, the

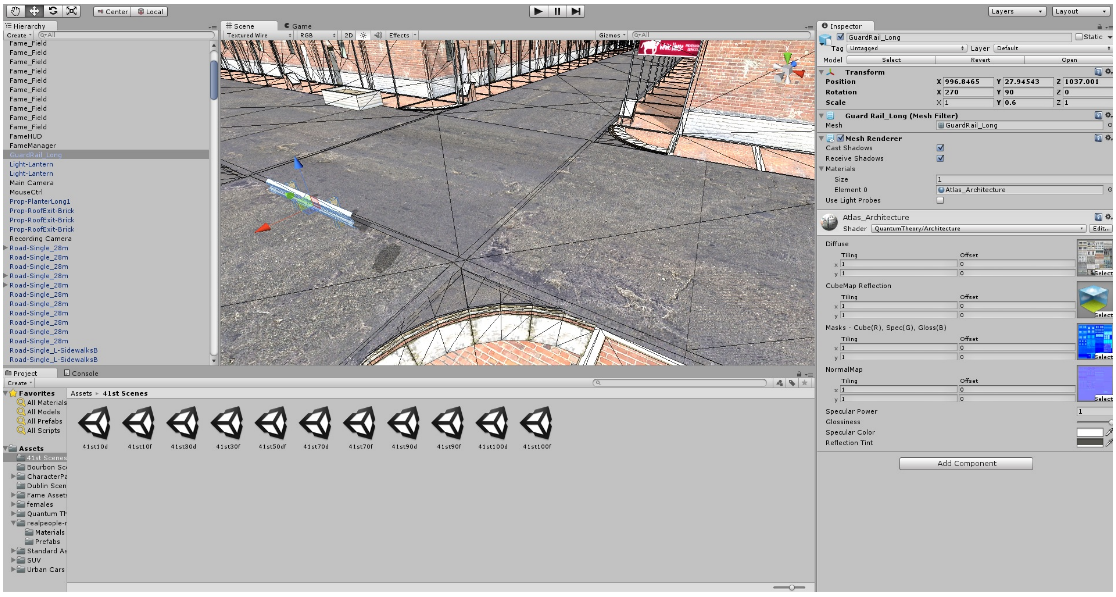

FIGURE 5.10: A screenshot of the Bourbon Street virtual environment being developed within Unity. Here the specific geometry that makes up the environment can be seen, with the window on the left highlighting some of the different prefabs that were defined and employed within the scene.

main paths were defined. When initialised, groups of agent would have a starting node and a destination node. The calculated path of nodes between these two points would then be iterated through for overall navigation. As many paths were created within the virtual environments as needed to provide one path for each group of agents up to the maximum calculated number of groups that specific location would have. This means in some simulation runtimes all paths would be in use due to the maximum number of groups being initialised, but in others only some of the paths would be used due to the reduced number of groups. While some paths might not be in use, they all remain the same over all runtimes of the urban crowd simulation.

By combining these three elements, namely the virtual environments, the vector fields and the paths, agents could be initialised for the purposes required for the grouping dynamics psychophysical experiment as presented in Section 6.4.

5.9 Grouping Dynamics

The grouping dynamics displayed in crowds are another key element when considering crowd behaviour. By viewing real-world crowds, it is possible to see that various groups make up the body of the crowd. This can be a various number of groups at once, each

FIGURE 5.11: An example of the modelled virtual version of Bourbon Street and the real-life location it was based upon for comparison.

with a different number of individuals present. The grouping dynamics behavioural feature looks at these specific elements with respect to crowd behaviour. The behavioural feature consists of two variables:

1. **Group Frequency**: The number of groups present within the crowd at a given time step.

2. **Group Density**: The number of individual agents present within each group or the overall group cardinality.

In order to implement this behavioural feature, the steering mechanics were updated. This consisted of various new steering forces to act as a control system for the groups.

FIGURE 5.12: An example of the modelled virtual version of 41st Street and the real-life location it was based upon for comparison.

A cohesion force dictates how closely each member of the group sticks together. An alignment force dictates how closely the group follows in formation. A separation force determines how far agents in the group can move apart from each other. A separation from other groups force, determines how the group as a whole will move to avoid colliding with other groups. These forces consist of two variables, a weight value to determine how much impact a specific force has in relation to the other forces and a radius to determine the size of the local neighbourhood considered for that specific force. Due to the fact that it is the grouping dynamics behavioural feature that is being altered for stimuli generation alone, the values for these steering mechanics are kept constant for all agents within the simulation, across all runtimes.

FIGURE 5.13: An example of the modelled virtual version of Temple Bar and the real-life location it was based upon for comparison.

For the grouping dynamics behavioural feature, an overall group control system was implemented. This allowed for the initialisation of a number of groups defined at runtime through an area calculated and visualised through a connected spline. Each group has a variable for the number of agents present, which can be altered as needed for generating stimuli. Once the number of agents in a group is defined, individual agents are selected from a large quantity of agent prefabs previously created. In this manner the majority of agents seen within the simulation will be unique, so as to not influence results by using only a few different agent models that are repeated obviously within the scene. The final positions of the agents within the group are dictated by the new steering mechanics, so as to ensure that the difference between the groups themselves are kept to a minimum. Each agent has a number of key variables, including a specific size represented as a radius, a maximum force value and a mass. Again, as the behavioural feature being tested is specifically related to the group frequency and density, these values are kept constant for all agents within the simulation. Figure 5.14 shows an example of the group control system in action through the initialisation of some agent groups.

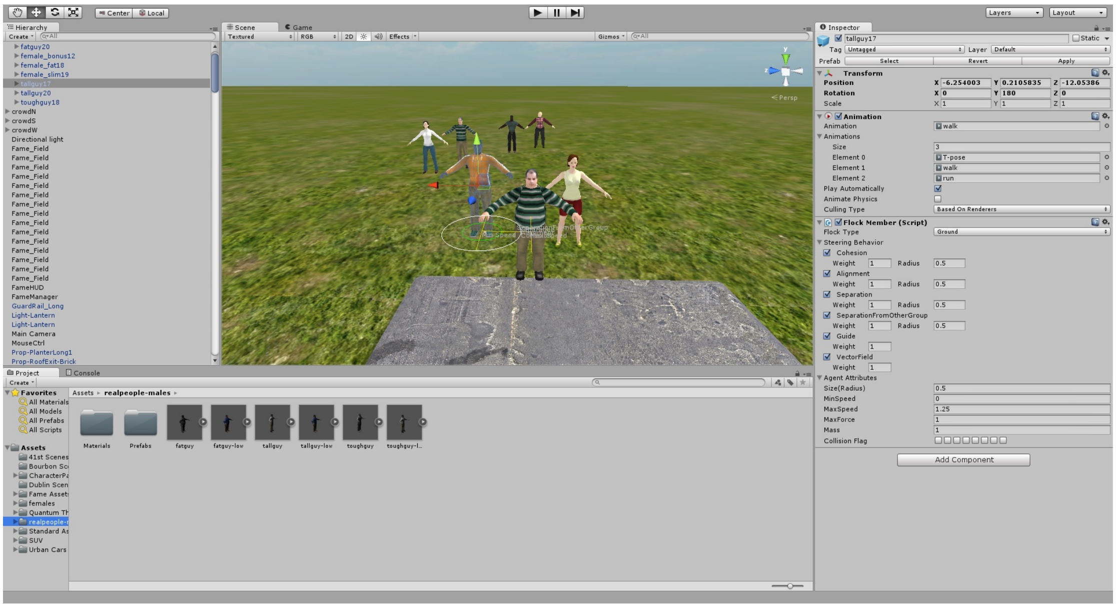

FIGURE 5.14: A screenshot showing the group control system with the initialisation of some agent groups. The window on the right highlights the various steering force parameters that are applied to the agents using this system.

The crowd grouping behavioural feature, doesn't communicate with the core algorithms in the same manner as the previous two features. It does however influence the core algorithms though the initialisation of the agents and groups. This has an overall effect on both the other algorithm, especially considering the new steering mechanics and the resulting crowd behaviour. Figure 5.15 shows this influence that grouping dynamics has on the core algorithms and crowd behaviour.

As noted, the grouping dynamics behavioural feature could be altered in terms of intensity for both group frequency and group density. The specifics of the intensity changes are dependent upon the crowd data for groups calculated at each real-world location. In order to generate different stimuli for psychophysical evaluation, the number of groups initialised and the number in each group can be specifically set before main simulation runtime. This involves using graphical interface elements, which were implemented

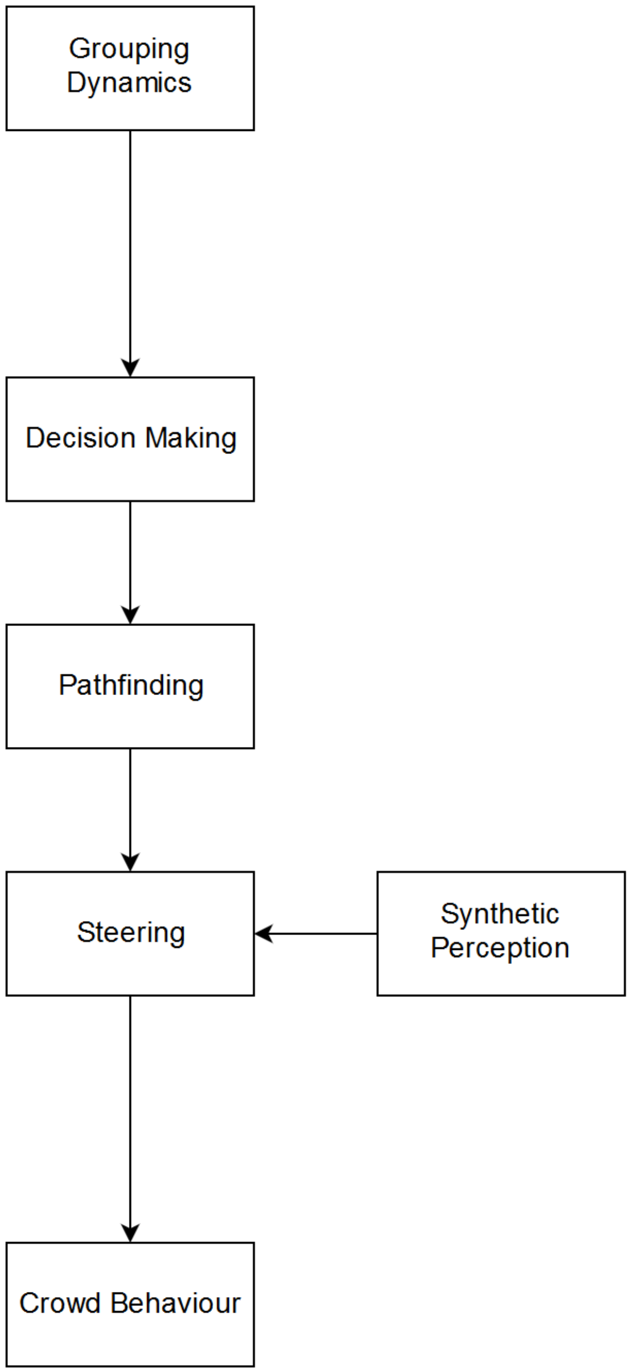

FIGURE 5.15: A diagram showing the influence of the grouping dynamics behavioural feature on the core algorithms and resulting crowd behaviour.

to make the generation of stimuli a more efficient process. To highlight this, Figures 5.16, 5.17 and 5.18 show a comparison at each location between a setup with low density crowds and one with high density crowds. The psychophysics experiment on the grouping dynamics behavioural feature is presented in Section 6.4.

FIGURE 5.16: Here is an example of the grouping dynamics behavioural feature at the Bourbon Street location. The top image shows the scene with low density crowds, whereas the bottom image shows the scene with high density crowds.

5.10 Summary of Chapter 5

In this chapter, the implementation of the urban crowd simulation has been examined. The need for a platform that can be refined over time through the iterative general methodology has been discussed, in addition to the behavioural features requiring the specific implementation of parameter spaces and customisability for producing stimuli for psychophysical evaluation. The development cycle of the simulation began with the core AI algorithm including finite-state machines for decision making, crowd path following

FIGURE 5.17: Here is an example of the grouping dynamics behavioural feature at the 41st Street location. The top image shows the scene with low density crowds, whereas the bottom image shows the scene with high density crowds.

for steering and A* for pathfinding and was then extended with additional features. The implementation of the behavioural features are outlined, with varying velocity, social forces and grouping dynamics. The parameters of each have been highlighted and it is noted how they can be varied to produce different intensities for stimuli generation. Finally, the overall refinement of the urban crowd simulation has been reviewed, with a focus on the development environments utilised, namely C++ OpenGL and then the Unity game engine. The reason for these choices and the advantages of each, have been noted.

FIGURE 5.18: Here is an example of the grouping dynamics behavioural feature at the Temple Bar location. The top image shows the scene with low density crowds, whereas the bottom image shows the scene with high density crowds.

In the next chapter, the three experiments conducted in this research are presented. The first experiment deals with the varying velocity behavioural feature, the second experiment covers the social forces behavioural feature and the third experiment outlines the grouping dynamics behavioural feature. The details of each are examined with relation to the feature itself, the creation of stimuli, apparatus and participants, procedure and results. Finally, a discussion is covered at the end of the chapter looking at the experiments as a whole.

Chapter 6

Experiments

In this chapter, the psychophysical experiments conducted as part of the perception stage of the general methodology are presented. This includes three experiments, each examining a specific behavioural feature. Experiment 1, covers the varying velocity feature, experiment 2, the social forces feature and experiment 3, the grouping dynamics feature. For each experiment details are outlined regarding the behavioural feature, creation of stimuli, apparatus and participants, procedure and results. Finally, a discussion regarding the experiments as whole is presented.

6.1 Purpose of Experiments

It has been shown that realism is an important aspect when considering simulation, particularly for virtual crowds. As the sophistication of algorithms and technology increases, so does the requirement of achieving realism. This is not straightforward however, as the sophistication of algorithms is no guarantee of incorporating realism and can potentially lead to the 'complexity fallacy'. For crowd simulation, the defining factor in terms of realism is typically the crowd behaviour exhibited by the agents. For achieving realism, it has been highlighted in the literature that both adherence to crowd data and the perceptual plausibility of the crowd behaviour is of concern. This virtual and perceived realism plays an important role and that is what these experiments explore, a means of evaluating crowd behaviour through this realism and perceptual plausibility.

To this end, these experiments apply the methods of psychophysics, a psychological approach for gauging human perception with respect to different intensities of stimuli. By using crowd data to inform initial stimulus creation, different intensities of stimuli can then be created through percentage-based alterations. In this manner, psychophysical methods are adapted and focused with behavioural features, in order to assess perceived realism while also considering virtual realism. The absolute thresholds are identified, along with the most perceptually realistic intensities. This means that the effectiveness of the crowd behaviour can be quantified to help future simulations achieve their purpose. These experiments are exploratory in nature by highlighting potential links between both virtual and perceived realism, opening avenues for further experimentation and new hypotheses.

6.2 Experiment 1: Varying Velocity

The first experiment presented in this thesis examines the varying velocity behavioural feature. The following Subsections will outline the behavioural feature and its variables, the testing apparatus and the number of participants. The experimental procedure is covered in detail and the results are discussed.

6.2.1 Varying Velocity

Varying velocity was the first behavioural feature that was identified as part of the analysis stage, as discussed in Section 4.1.1. By looking at reality, it is easy to see crowds and their composite components moving at various speeds. This may initially seem like an obvious observation, but it is a factor often not fully explored when considering implementing virtual crowds. There is space for questions with regards to velocity magnitudes; ‘What are effective minimum and maximum velocities?’ ‘Is there a specific velocity range that is the most plausible?’ ‘What is the best distribution, velocities closer to the minimum or maximum?’ Indeed, by studying the velocity of individuals and groups within a crowd, it is possible to see constant fluctuation. In addition, velocity is the main movement force applied to agents for simulation (Fang et al. 2012) and varying its properties can have a noticeable impact upon the resulting behaviour. Both the prominence of velocity as an element of crowds and the wide applicability of velocity for simulating crowds, highlight the benefits of evaluating it as a behavioural feature.

When considering the velocity with respect to crowds, two main variables can be identified:

- **Velocity Range**: This variable determines the minimum and maximum velocity magnitudes available to the agents, in essence the different speeds they can operate within.
- **Velocity Distribution**: This variable determines the distribution for the average magnitude utilised within the velocity range. When referring to agents, this is often called the preferred velocity and is essentially the speed that agents will try to maintain.

By varying the range and distribution with respect to the magnitude of agent velocity, the resultant crowd behaviour can be altered to a degree that is visible to viewers. This directly affects the overall perceptual plausibility of the implementation, which can then be measured as perceived realism through psychophysical evaluation. This leads to the identification of perceptual metrics, such as thresholds for plausibility and optimum configurations. The results can be used for implementations of virtual crowds and as reference for related future work.

6.2.1.1 Psychophysical Method

The psychophysical method selected to assess the varying velocity behavioural features was a staircase procedure, which falls into the category of an adaptive method. Additional elements were applied to further enhance the base method, such as using a randomly interleaved approach with ascending and descending staircases in a Three-Down One-Up (3D1U) configuration, using a total of four staircases per participant. Specifics regarding the procedure of applying this method as part of this experiment are discussed in Section 6.2.4.

The main reasons for choosing a staircase procedure for this initial experiment is due to some of the advantages inherent in using an adaptive method (Leek 2001). Unlike its classical counterparts such as the method of limits, an adaptive method such as a staircase procedure, tailors what stimuli is presented based on a participants responses at each trial. In this manner, most of the data is collected from around what is termed the reversal point, or simply the absolute threshold. As determining the perceptual metrics for this behavioural feature is one of the main objectives for this experiment, the directed focus at the thresholds is important for improving accuracy. Given that at this stage of the research the psychophysics platform for experimentation was not yet deployed online, a smaller scale study was chosen and thus the benefits of improved accuracy and more data with respect to the absolute threshold were a major contributing factor in the choice of using a staircase procedure. Using a staircase procedure allows for the plotting of the psychometric function, again an important consideration for identifying perceptual metrics, however, one disadvantage of the adaptive method's focus is that it means less data is collected over the entire psychometric function. In the case for this smaller scale study however, the benefit to the focus outweighs this factor. Again, with this experiment being on a small scale, one of the disadvantages of staircase procedures being complex to carry out was controlled to a degree.

Additional elements were added to this staircase procedure such as to use randomly interleaved ascending and descending staircases, due to the fact it helps to eliminate an inherent disadvantage in the staircase procedure design, which is that complexity is needed or else it is possible for participants to guess the intensity steps. By using multiple interleaved staircases, it makes participants being able guess the scale based on their choices at each trial far harder than with just one staircase. In addition, the random factor in the interleaving helps this by adding an element of the widely-used method of constant stimuli. Similarly, by employing a 3D1U staircase type that requires multiple correct responses to indicate a reversal and adaptive step size that reduces after some reversals, it improves the overall accuracy of the collected data and controls elements that can cause bias in standard procedures (Garcıa-Pérez 1998).

6.2.2 Stimuli Creation

A number of short video clips were required as stimuli for psychophysical evaluation of the varying velocity behavioural feature. An example of the final stimuli can be seen in Figure 6.1.

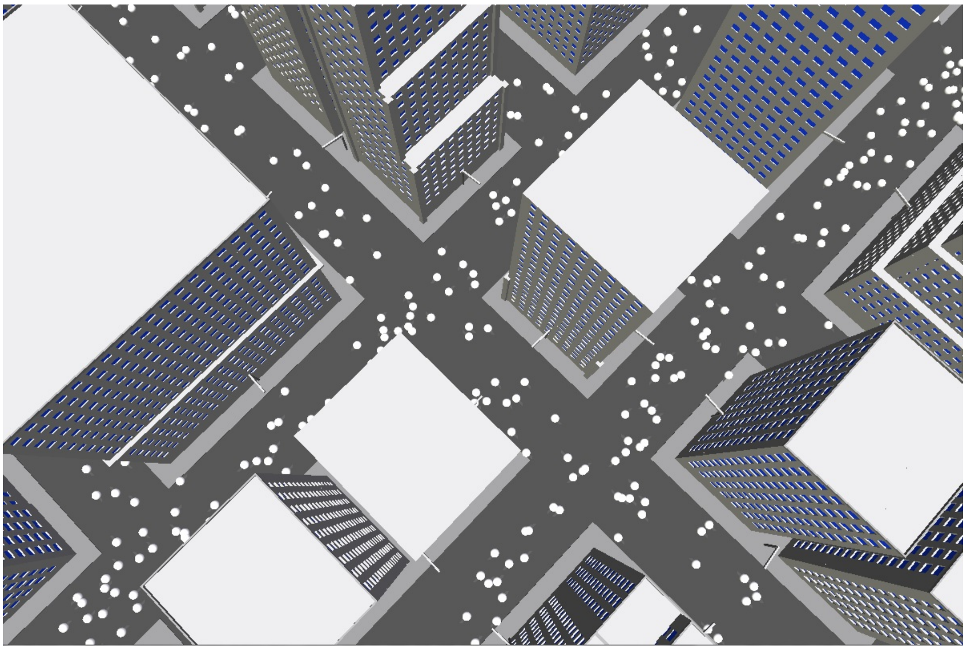

Figure 6.1: An example of the video stimuli recorded for the varying velocity experiment.

The implementation details regarding this behavioural feature were outlined in Section 5.6, however, to summarise, by altering two parameters, agent velocity can be restricted to certain limits with respect to magnitude range and distribution. To create a stimulus, the specific parameter values required are passed and the simulation is then initialised. Once at runtime, video capture software is used to record a video clip for approximately ten seconds showcasing the virtual crowds. The process is repeated using different parameter values, in order to set different intensities with respect to the behavioural feature. For this experiment a total of fourty video clips were recorded at 60 frames per second as stimuli, each with a different intensity for velocity range and magnitude. Aside from the two variables being altered, all other parameters and algorithms were kept constant during stimuli creation. This ensures that the only visible differences in terms of the crowd behaviour, are due the alteration in the intensity of the behavioural feature.

6.2.2.1 Crowd Data

Crowd data was analysed to ensure that the stimuli was initially tempered and grounded within reality, so that any resultant data would be applicable towards various types of simulation. With the varying velocity behavioural feature as the focus for this analysis, a manual annotation was tailored for examining the various velocities displayed by a selection of pedestrians over multiple locations. These locations, as can be seen in Figure 6.2, were:

1. Times Square in New York.
2. Bourbon Street in New Orleans.
3. Abbey Road in London.

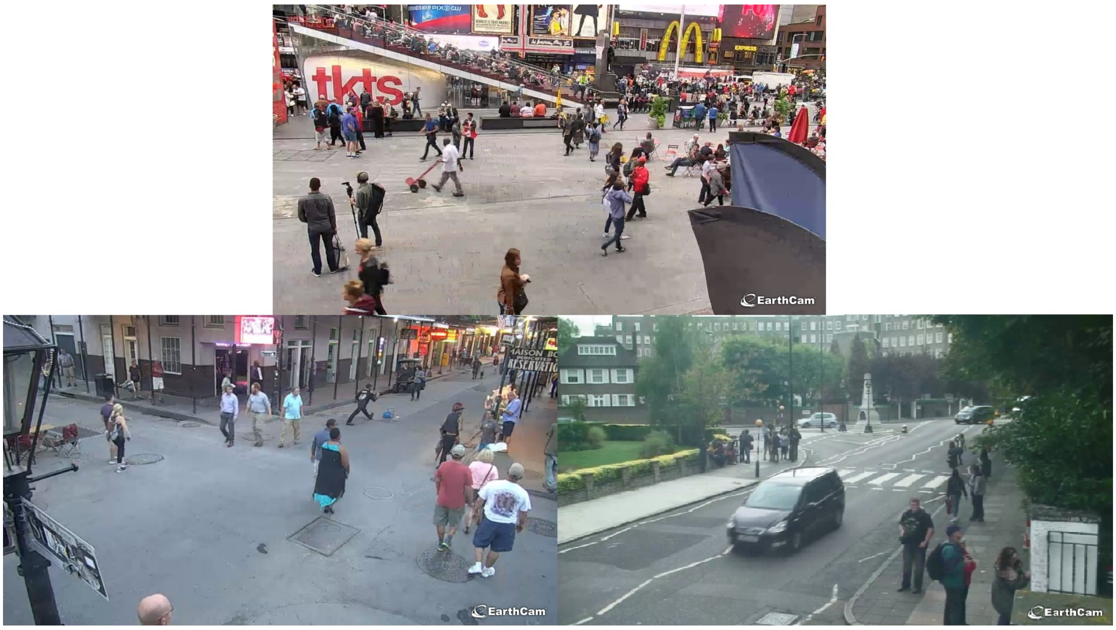

FIGURE 6.2: These three locations provided data with regards to the varying velocity behavioural feature. Top-middle shows Times Square, bottom-left, Bourbn Street and bottom right, Abbey Road.

The camera position at these locations, along with the clarity and volume of pedestrian traffic made them well suited for examining velocities. By employing visual identification, a selection of pedestrians were annotated with respect to their distance travelled over small timeframes. Figure 6.3 shows an example of this process, how a pedestrian is tagged and their distance is highlighted. Using these distances over time, the average velocity magnitude could be calculated for all pedestrians. The resulting value was:

- Average Velocity Magnitude: 1.25 m/s.

Preferred pedestrian walking speed is typically determined at 1.4 m/s (Browning et al. 2006), showing the analysed result falls close to this figure albeit slightly lower.

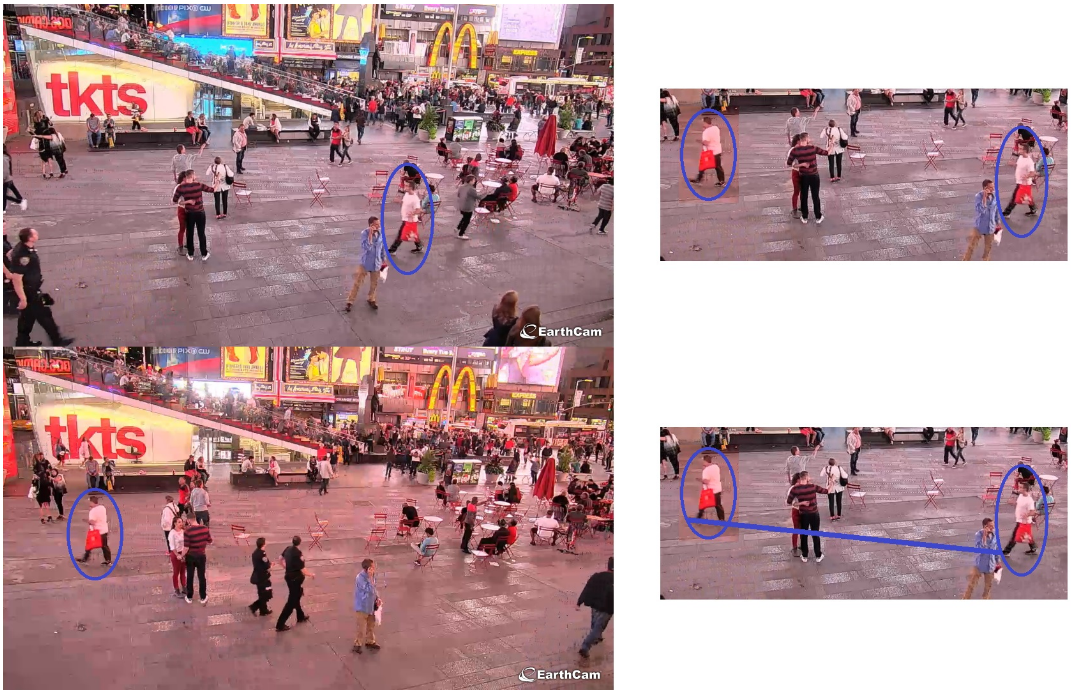

FIGURE 6.3: A key frame from one of the selected video clips. This shows an example of the annotation process whereby a pedestrian is tagged over a small timeframe and the distance is highlighted.

6.2.2.2 Variation

Psychophysical experiments require different intensities with regards to either a stimulus or stimuli. By employing the three-stage general methodology and analysing crowd data, the initial value for average velocity magnitude was obtained. This value was then set as 50% intensity for the purpose of varying both up and down the scale. This scale was percentage based, as is typical for psychophysical studies as it allows for a wide range of stimuli to explore the psychometric function. Table 6.1 shows the results of this process, with the different intensities and what they relate to in terms of magnitudes for both velocity range and distribution. These values link directly to the velocities displayed by the agents in meters per second. Note, for staircase procedures additional step sizes can be used to explore the area around thresholds, hence a larger pool of stimuli typically created compared to other psychophysical methods.

6.2.3 Apparatus and Participants

A total of three participants (2 males and 1 females, aged between 18 and 35) from various educational backgrounds but all with basic computing skills, took part in the

Intensity (%)	Range (m/s)	Lower (m/s)	Upper (m/s)	Distribution (m/s)
5	0.125	1.1875	1.3125	0.125
10	0.250	1.1250	1.3750	0.250
15	0.375	1.0625	1.4375	0.375
20	0.500	1.0000	1.5000	0.500
25	0.625	0.9375	1.5625	0.625
30	0.750	0.8750	1.6250	0.750
35	0.875	0.8125	1.6875	0.875
40	1.000	0.7500	1.7500	1.000
45	1.125	0.6875	1.8125	1.125
50	1.250	0.6250	1.8750	1.250
55	1.375	0.5625	1.9375	1.375
60	1.500	0.5000	2.0000	1.500
65	1.625	0.4375	2.0625	1.625
70	1.750	0.3750	2.1250	1.750
75	1.875	0.3125	2.1875	1.875
80	2.000	0.2500	2.2500	2.000
85	2.125	0.1875	2.3125	2.125
90	2.250	0.1250	2.3750	2.250
95	2.375	0.0625	2.4375	2.375
100	2.500	0.0000	2.5000	2.500

TABLE 6.1: The various intensities linked to the specific range and distributions for velocity magnitude.

first experiment. All participants were naïve as to the overall purposes of the experiment, but were informed as to what they were judging on each trial. Due to the fact that the main two variables being considered were in relation to crowd behaviour in terms of varying velocity, the other factors that could influence crowd behaviour were kept the same between stimuli.

A specially developed psychophysics platform was used for this experiment. Figure 6.4 shows a page from this platform. A participant was presented with instructions and had to agree to be part of the experiment to continue. Next some basic demographic data was taken and the experiment began. A series of trials were then presented to each participant.

6.2.4 Procedure

This experiment examines the varying velocity behavioural feature with two variables, velocity range and velocity distribution. The psychophysical method employed was a

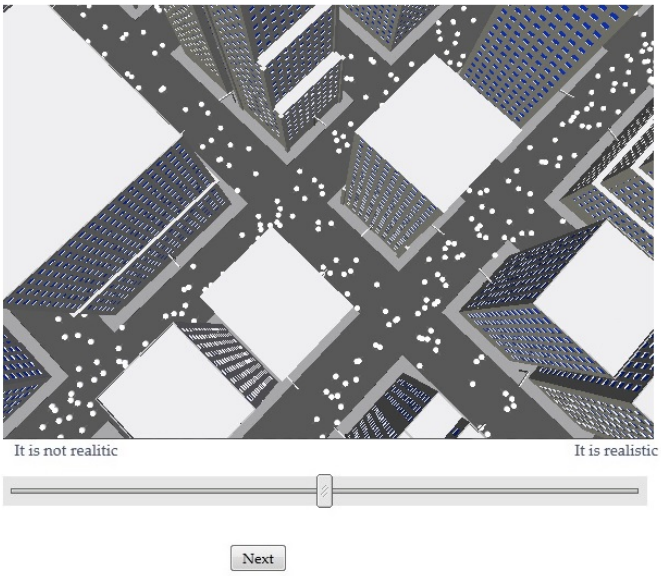

Figure 6.4: A page from the experiment platform showing a trial. The stimulus contains a specific intensity for a variable of the varying velocity feature.

staircase procedure, using a randomly interleaved approach with both ascending and descending staircases. A staircase procedure was chosen as the method allows for responses to be kept within proximity of thresholds, an important consideration for experiments with only a small pool of participants. By randomly interleaving the staircases it ensures that participants cannot guess the scale, an issue present with some basic staircase procedures. Using both ascending and descending staircases, allows for both ends of the psychometric function to be considered.

Each variable was analysed using real-world crowd footage to provide a starting point for stimuli, as outlined in Subsecton 6.2.2.1. For the two variables, a total of twenty intensities were selected. This allowed for enough stimuli to probe the thresholds, while also keeping the experiments length within reason so as to not dissuaded participants. These intensities were:

- 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% and 100%.

By employing a staircase procedure, it means the presented stimuli intensities are based upon participants responses. Initial step-size was set at 10% to allow for efficient identification of the threshold area, however the step-size was reduced to 5% after the 4^{th}

reversal. 3D1U type of staircase procedure was utilised, meaning that three seen responses were required for an intensity reduction and one unseen response for an increase. A reversal occurs when the observer's response changes from one response to the other, for example from a seen response to an unseen response. A total of eight reversals occur before a staircase is complete. A total of four staircases were used, two for each variable. All staircases started at 50% intensity, the configuration extracted from the crowd data, however one per variable would be increased in intensity initially and the other decreased. Employing an adapted step-size 3D1U staircase type, with ascending and descending staircases randomly interleaved, improves the overall accuracy of the collected data and controls elements that can cause bias in standard procedures (Garcıa-Pérez 1998).

The stimuli itself was forty recorded video clips of approximately ten second duration, as noted in Subsection 6.2.2. The different stimuli intensities were created through varying the velocity magnitude and range, which were calculated from the crowd data. This was discussed in Subsubsection 6.2.2.2, with Table 6.1 mapping the different intensities to their specific values. It should be noted that the stimuli for 50% intensity is always based on the values from the crowd data, providing a focal point between reality and perception.

A local platform was used to present the experiment to participants, as outlined in Subsection 6.2.3 and shown in Figure 6.4. Each intensity presented to participants is called a trial and on each trial the participants evaluate the virtual footage. They select a percentage value from 0% to 100%, depending on how realistic they find the crowd behaviour. Higher percentages determine it to be more realistic and lower percentages mean the participants found it unrealistic. A percentage value of 50% or higher is considered a seen-response for the purpose of reversals, psychometrics and threshold calculation. In addition, this percentage value is also converted to a perceived realism value for finding the highest response and identifying the optimum configurations in terms of the behavioural feature.

The velocity range trials were presented first and the velocity distribution trials followed. On each trial a participant is shown the simulated footage with a specific intensity for the behavioural features variables. They were asked about the realism of the displayed crowd behaviour and had to select a percentage using a slider with a displaying value, before they could submit and proceed on. Note, the participants were informed as to the type of trial, whether it be velocity range or velocity distribution.

6.2.5 Results

The purpose of these results was to highlight the absolute thresholds for when each variable becomes perceptually realistic and identify the optimum configurations within these ranges. For a staircase procedure without any forced choice element, the absolute thresholds are taken at the 50% seen level. A seen result is classified as a perceived realism response of 50% or higher. For each variable, at each intensity, the number of seen results were collected and a percentage for the total number of seen results was calculated based on the pool of participants. By using psychometrics in conjunction with perceived realism, the highest average responses that fall within the perceptual thresholds can be identified.

As part of a staircase procedure, each participant produces a staircase of seen results. In this case, this includes both ascending and descending staircases using the 3D1U method for a total of eight reversals. An example participant ascending staircase for velocity distribution can be seen in Figure 6.5, showing reversals represented by the peaks and troughs. Being an ascending staircase, this Figure shows that the stimuli intensity was increased until it became unrealistic, at which point a reversal occurred until it became realistic again. As can be seen, a total of eight reversals occurred in this fashion to allow for psychometric data with respect to that participant to be collected. All participants produced both ascending and descending staircases via their responses in the experiment, which gave targeted psychometric data around the thresholds.

The psychometric function was plotted using the psychometric data from the participant staircases for both the velocity range and velocity distribution variables, as shown in Figures 6.6 and 6.7 respectively. For both graphs, percentage seen responses are on the ordinate and stimuli intensity on the abscissa, as is typical in psychophysics. The overall shape of these psychometric functions were as expected, with a peak in terms of the percentage seen responses that drops off as the stimuli intensity is either increased or decreased away from this optimum level. In addition, both graphs show that the range of stimuli intensities were broad enough to have allowed for the absolute thresholds to be identified, due to the point at which the psychometric function crosses 50% seen on the ordinate for both the upper and lower thresholds being visible. Note, that both graphs begin and end on slightly different stimuli intensities due to the data points collected as part of the staircase procedure. As the staircase procedure is adaptive, the intensity of stimuli presented depended upon the responses and thus some intensities never required probing.

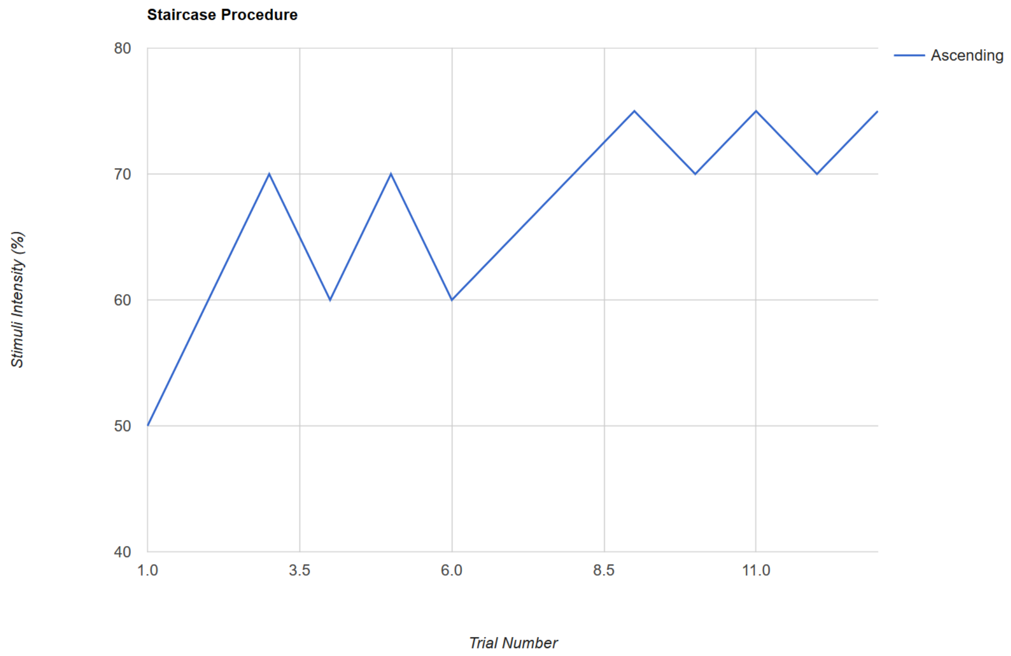

FIGURE 6.5: A graph showing an ascending staircase with regards to velocity distribution for a single participant. The peaks and troughs show a total of eight reversals.

As can be seen in Figure 6.6, the absolute thresholds for velocity range were identified as 19% intensity and 51% intensity, which links to velocity ranges of 0.4875m/s and 1.2625m/s. By keeping the velocity range at or within these parameters, the crowd behaviour with respect to this element remains perceptually plausible. See Table 6.1 to see the specific upper and lower range these values correspond with. As can been seen in Figure 6.7, the absolute thresholds for velocity distribution were identified as 29% intensity and 70% intensity, which links to velocity distributions of 0.725m/s and 1.75m/s. By keeping the velocity distribution at or within these parameters, the crowd behaviour with respect to this element remains perceptually plausible.

The optimum configuration was identified using the perceived realism values collected through the psychometric data. As can be seen in Figure 6.8, for velocity range, the highest perceived realism value of 0.82 can be seen at 30% intensity, equating to a value of 0.75m/s. As can be seen in Figure 6.9, for velocity distribution, the highest perceived

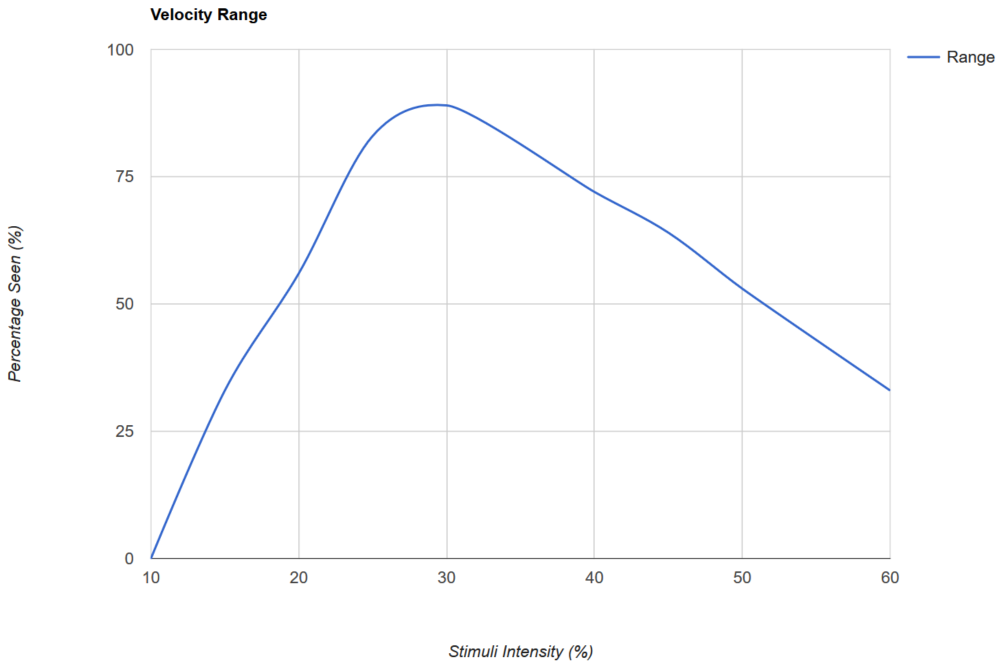

FIGURE 6.6: A graph showing the psychometric function for velocity range, with percentage seen responses on the y-axis and stimuli intensity on the x-axis.

Varying Velocity	Lower Threshold	Upper Threshold	Optimum
Range	0.4875 m/s	1.2625 m/s	0.75 m/s
Distribution	0.725 m/s	1.75 m/s	1.25 m/s

TABLE 6.2: A summary of the absolute thresholds and optimum configuration for the varying velocity behavioural feature.

realism value of 0.85 can be seen at 50% intensity, equating to a value of 1.25m/s. For both graphs, a missing bar signifies that no data was collected for that specific intensity of stimuli. This occurred due to the participant responses influencing what intensities were shown as part of the staircase procedure, meaning in those cases the intensity did not need to be presented. This is also the reason why both graphs start and end at different stimuli intensities. A summary of absolute thresholds and optimum configuration are shown in Table 6.2. Further discussion and the difference between virtual and perceived realism are presented in Section 6.5.

6.3 Experiment 2: Social Forces

The second experiment presented in this thesis examines the social behavioural feature. The following Subsections will outline the behavioural feature and its variables, the

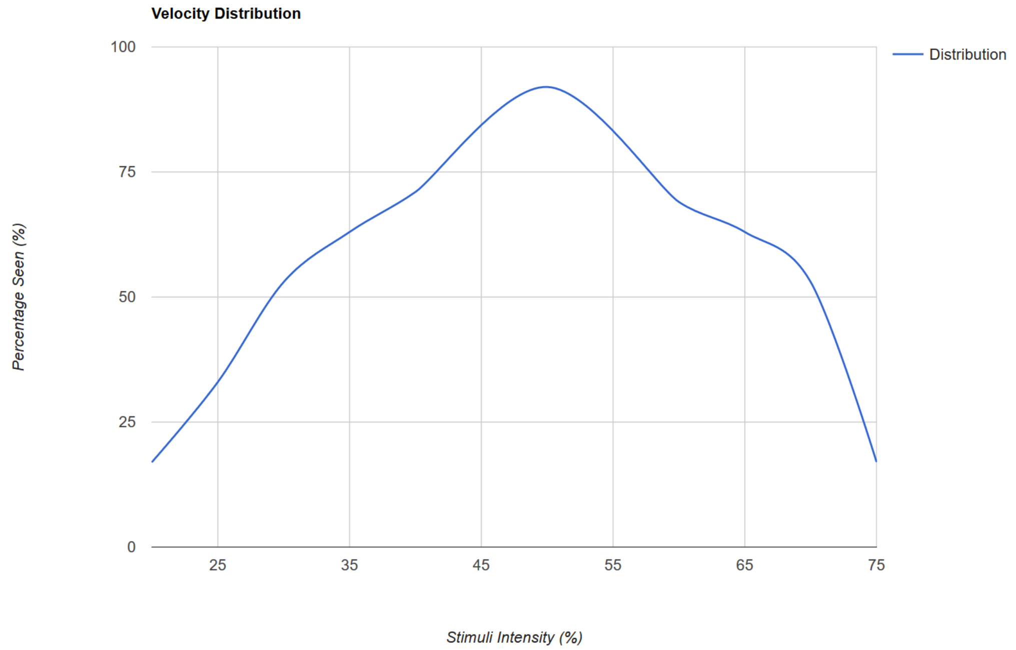

FIGURE 6.7: A graph showing the psychometric function for velocity distribution, with percentage seen responses on the y-axis and stimuli intensity on the x-axis.

testing apparatus and the number of participants. The experimental procedure is covered in detail and the results are discussed.

6.3.1 Social Forces

For the second iteration of the analysis stage, the behavioural feature for social forces was identified as part of the analysis stage, as discussed in Section 4.1.1. The purpose of social forces is specifically to improve crowd behaviour by describing the internal motivations present within each member of a crowd. Social forces are prominent for implementation when realistic crowd behaviour is required, as seen in the literature (Almeida et al. 2013, Szymanezyk et al. 2012, Mehran et al. 2009). The effects of social forces are seen within real-world crowds when certain emergent behaviours occur. For example, some individuals form groups while some keep out of close proximity with others. It can be said that social forces, especially those concerning the interactions between agents, are a core element of crowd behaviour, which are easily seen through their results. These considerations show social forces to be a prime behavioural feature for psychophysical experimentation.

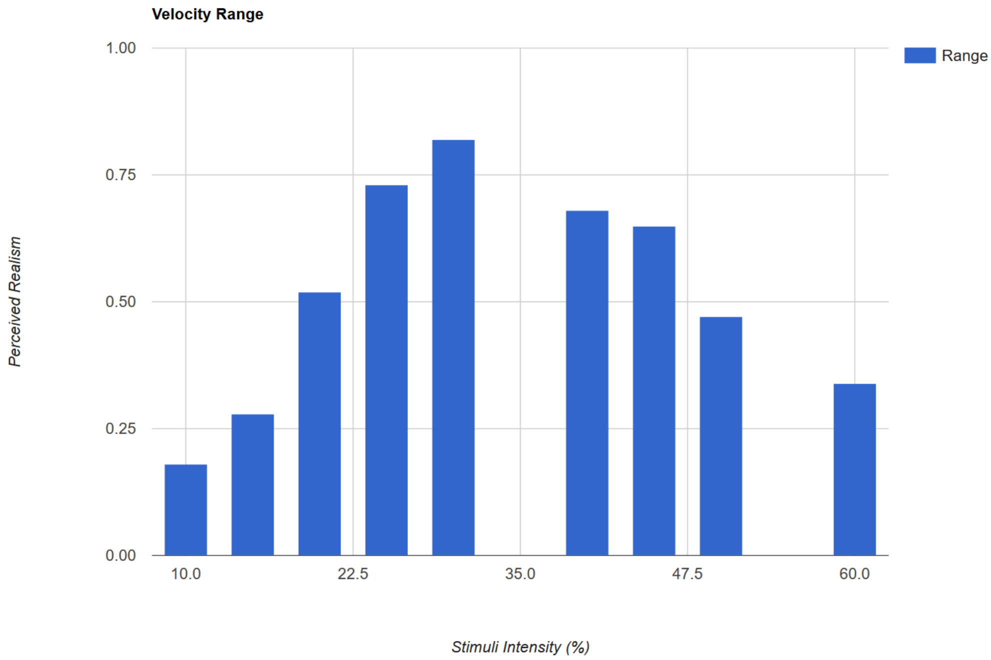

FIGURE 6.8: A graph showing the perceived realism values for velocity range.

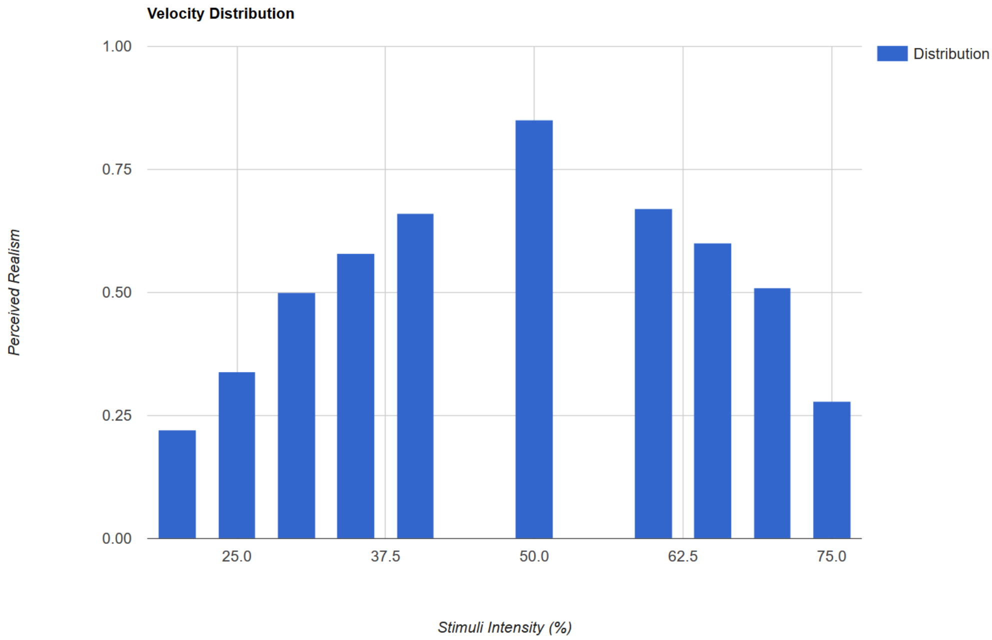

FIGURE 6.9: A graph showing the perceived realism values for velocity distribution.

When examining the social forces between individual agents within a crowd, two main variables can be seen as outlined in Helbing's model (Helbing & Molnar 1995):

- **Agent Repulsion**: The force of repulsion between agents, which causes certain agents to move away from other agents. This force causes agents to stay out of proximity with one and another, leading to more individualistic behaviour.

- **Agent Attraction**: The force of attraction between agents, which causes certain agents to move towards other agents. This force causes agents to move closer together and form more ordered group behaviour.

By altering these two forces the emergent crowd behaviour between agents can be altered to noticable degree, which in turn impacts the crowd's perceptual plausibility. By examining these social forces through psychophysical evaluation, perceptions of the crowd behaviour can be quantified by measuring perceived realism with regards to different intensities of each force. This allows for the identification of perceptual thresholds and optimum configurations, which can be of use for future implementations of social forces and future research.

6.3.1.1 Psychophysical Method

The psychophysical method selected to assess the social forces behavioural features was the method of constant stimuli, which falls into the category of a classical method. Additional elements were applied to further enhance the base method, such as using a 2AFC approach. Specifics regarding the procedure of applying this method as part of this experiment are discussed in Section 6.3.4.

The main reasons for choosing the method of constant stimuli for this second experiment is due to both the further refinement of the urban crowd simulation and the update to add the psychophysical platform online to have a larger scale study. For the previous smaller scale experiment with varying velocity, certain aspects of the adaptive method were advantageous such as focusing data around the threshold, however this can also be a limitation when it is the first method applied to a behavioural feature. Probing across the whole psychometric function or intensity range has its advantages, especially considering in these experiments both upper and lower absolute threshold are considered, along with finding the optimum intensity that is typically not close either of the thresholds. As such, a classical approach allowed for this wide data acquisition and was significantly more straightforward to implement as part of the online platform. In addition, the method of constant stimuli is particularly effective (Simpson 1988) and perhaps the most widely utilised classical psychophysical method owing in part to its inherent element of randomness, which helps to migrate the chance of participants identifying the intensity scale. Constant stimuli as part of its design allows for the both psychometric function to be plotted and the absolute thresholds to be identified, which are major considerations in these experiments.

The additional element of incorporating 2AFC to the method of constant stimuli was due to some of the benefits it can provide for experimentation (Shelton & Scarrow 1984). By imposing a time limit and giving options between videos with the behavioural feature present and those without, further data is gained with respect to differentiation and in many cases efficiency can be increased. In addition, in some cases the objectivity of the participants is increased and by providing comparative stimuli on each trial, participants typically have higher responses than in experiments without forced choice. This and adjusting for the chance of participants guessing due to the forced choice, is the reason why for forced choice experiments it is common to take the absolute threshold at the 75% seen level rather than the more standard 50%. Forced choice has been utilised successfully for testing the discrimination of motion perception (Gold & Shadlen 2000), so given the type of stimuli utilised in this experiment, the application was highly suited for this specific instance.

6.3.2 Stimuli Creation

A number of short video clips were required as stimuli for psychophysical evaluation of the social forces behavioural feature. An example of the final stimuli can be seen in Figure 6.10.

The implementation details regarding the social forces within the urban crowd simulation were outlined in Section 5.7. To summarise, by altering the weight values of the specific social forces, the prominence of each force can be controlled. Note, three forces are present within the implementation, however only the agent forces are varied, whereas the object repulsion force is kept constant to support collision detection with scene geometry. Once the correct weights are set for the forces, the simulation is initialised. Once runtime begins, video capture software records a video clip for approximately ten seconds. The process is then repeated with different weight values for the behavioural feature. For this experiment a total of eighteen video clips were recorded at 60 frames per second as stimuli, each with a different intensity of the social forces. In addition,

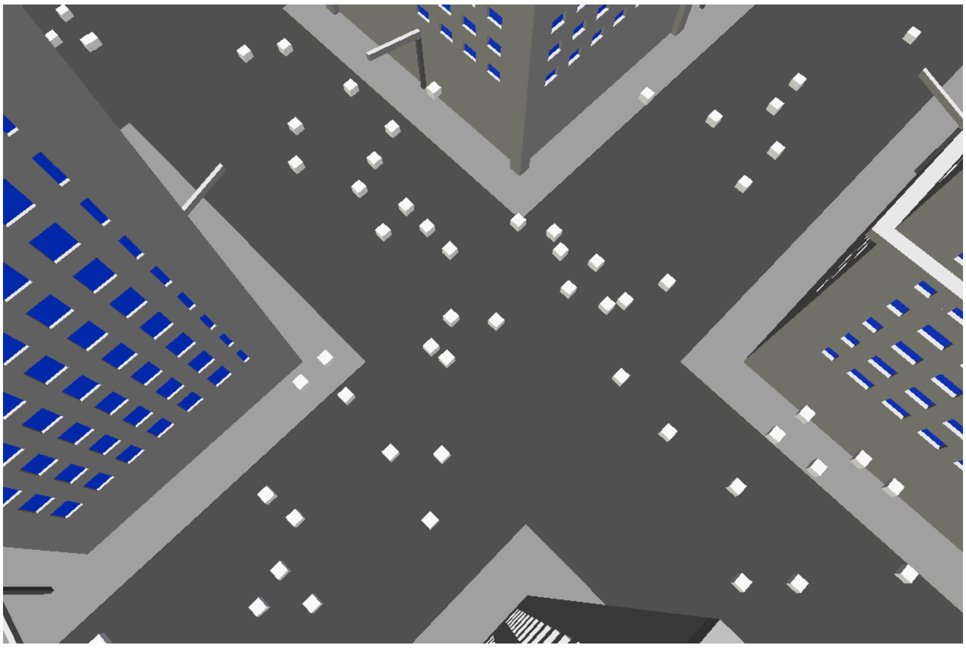

FIGURE 6.10: An example of the video stimuli recorded for the social forces experiment.

eighteen extra videos, each with a duration of ten second showing a control without the behavioural feature present, were recorded. Aside from the two forces being altered, all other parameters and algorithms were kept constant during stimuli creation. This ensures that the only visible differences in terms of the crowd behaviour are due the alteration in the intensity of the social forces.

6.3.2.1 Crowd Data

By analysing crowd data, the creation of the stimuli was grounded in reality and tempered to ensure results would be applicable to a wide range of virtual crowd implementations. With the social forces behavioural feature as the focus for analysis, a manual annotation was tailored for examining the attraction and repulsion forces between pedestrians at different locations. These locations, as can be seen in Figure 6.11, were:

1. 5th Avenue in New York.
2. Abbey Road in London.
3. Temple Bar in Dublin.

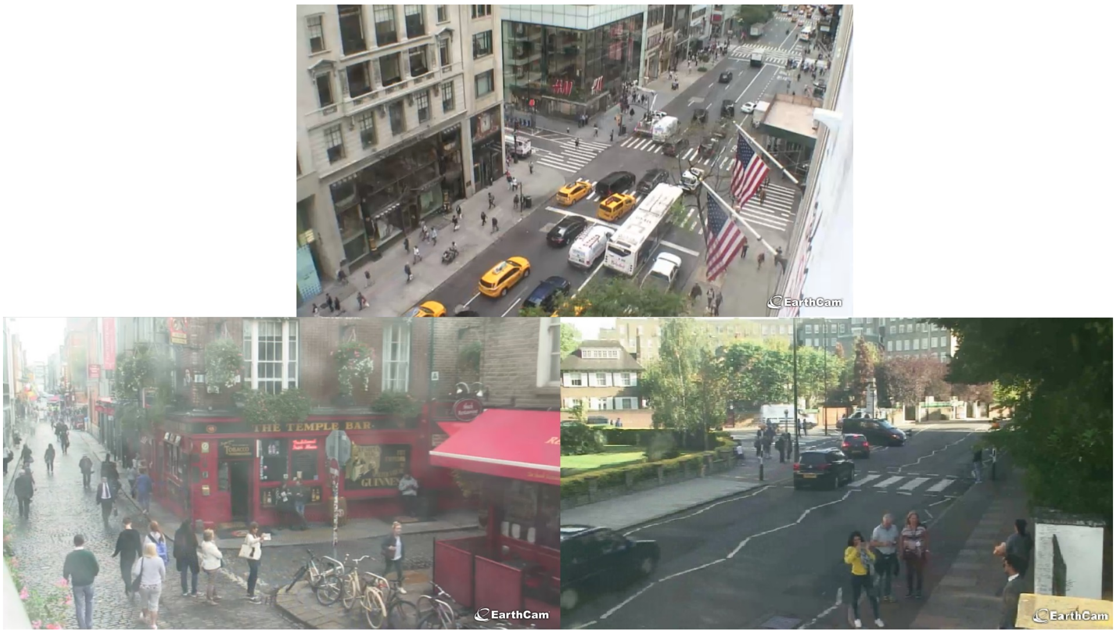

Figure 6.11: These three locations provided data with regards to the social forces behavioural feature.

These locations were ideal as they showed pedestrianised streets, with the camera location and visibility to accurately identify occurrences of these forces. Through visual identification at key frames over a series of intervals, a selection of pedestrians shown to display these behaviours were annotated and the overall frequency of the forces recorded. Figure 6.12 shows an example of this process, where forces of attraction and repulsion are highlighted. By annotating the pedestrians with respect to these forces, it was possible to find a similarity between force weight and displayed frequency for reproduction in simulation. The resulting values were:

- Approximately 50% of Pedestrians Displayed the Agent Attraction Force giving a Weight (w) value of 0.5.
- Approximately 50% of Pedestrians Displayed the Agent Repulsion Force giving a Weight (w) value of 0.5.

6.3.2.2 Variation

Psychophysical experiments required a range of different stimuli that are produced by varying intensity. By analysing crowd data, initial weights for the different social forces between agents were identified. These weight values were then set as 50% intensity, so

FIGURE 6.12: A key frame from one of the selected video clips. This shows an example of the annotation process whereby pedestrians are tagged and the social forces exhibited highlighted. The colour of the annotation denotes type of force being seen, blue for agent attraction and red for agent repulsion.

Intensity (%)	Agent Attraction (w)	Agent Repulsion (w)	Approximate Effect due to Weight (w) and Fluctuation Factor (f)
10	0.1	0.1	10% of Agents visibly display the force
20	0.2	0.2	20% of Agents visibly display the force
30	0.3	0.3	30% of Agents visibly display the force
40	0.4	0.4	40% of Agents visibly display the force
50	0.5	0.5	50% of Agents visibly display the force
60	0.6	0.6	60% of Agents visibly display the force
70	0.7	0.7	70% of Agents visibly display the force
80	0.8	0.8	80% of Agents visibly display the force
90	0.9	0.9	90% of Agents visibly display the force

TABLE 6.3: The different intensities linked to the weight values for the two forces as part of a social forces algorithm.

that both higher and lower intensity stimuli could be created. Variation was calculated through the percentages to provide a wide range of stimuli for exploring the psychometric function. Table 6.3 highlights the results of this process and shows the different intensities and what they relate to in terms of the two forces. These values link directly to the social forces displayed by the agents. For example, a weight value of 0.8 for agent attraction would translate to more agents forming groups and moving as units compared to a weight of 0.4.

6.3.3 Apparatus and Participants

A total of thirty-two participants (26 males and 6 females, aged between 18 and 35) from various educational backgrounds but all with basic computing skills, took part in this second experiment. All participants were naïve as to the overall purposes of the experiment, but were informed as to what they were judging on each trial. Due to the fact that the main two variables being considered were in relation to crowd behaviour in terms of social forces, the other factors that could influence crowd behaviour were kept the same between stimuli.

In order to reach a larger number of participants, a specially developed psychophysics platform was deployed online to improve accessibility. Figure 6.13 shows a webpage from this platform. Upon navigating to the webpage and beginning the test, a participant is presented with instructions and must agree to be part of the experiment to continue. Next some basic demographic data is taken and the experiment begins. A series of trials are then presented to each participant.

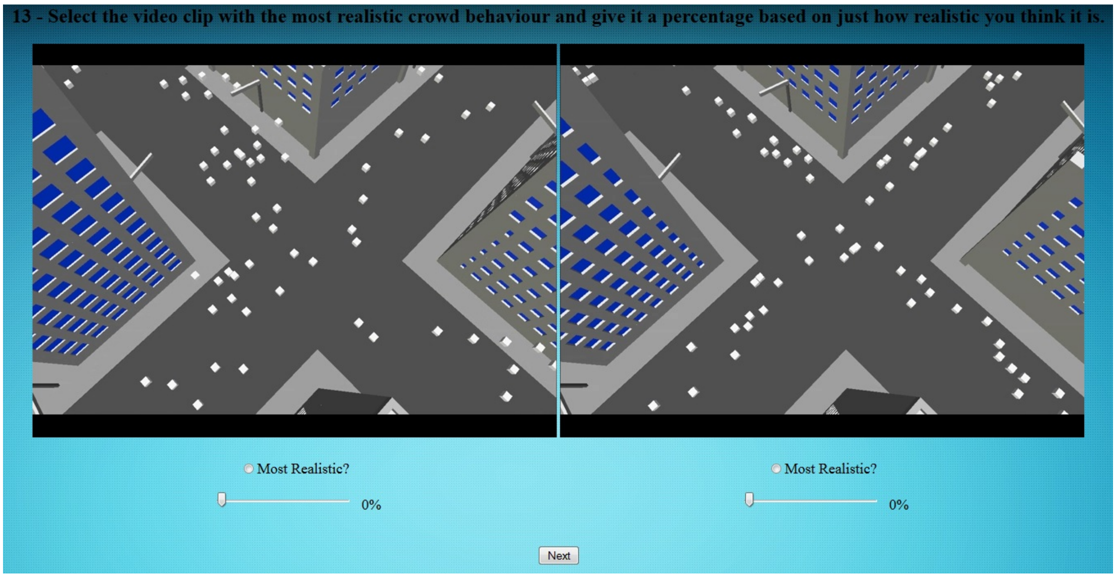

FIGURE 6.13: A webpage from the online platform showing a trial. The stimulus on the left contains a specific intensity of a social force and the other on the right a control with no social force present.

6.3.4 Procedure

This experiment examines the social forces behavioural feature with two forces, agent attraction and agent repulsion. The psychophysical method employed was the 2AFC approach, with constant stimuli. 2AFC was an appropriate method as it not only allows for wide probing of the psychometric function, but also has shown positive results for the discrimination of motion perception. The constant stimuli element controlled the potential of participants identifying the intensity scale through randomisation of stimuli order.

Each force was analysed using real-world crowd footage to provide a starting point for stimuli, as outlined in Subsecton 6.3.2.1. For the two forces, a total of nine percentage-based intensities were selected. This allowed for wide probing of psychometric function, while also keeping the experiments length within reason so as to not dissuaded participants. These intensities were:

- 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, and 90%.

For having an element of constant stimuli, a randomised order for presenting the different intensities was generated:

- Agent Repulsion: 70%, 20%, 90%, 60%, 30%, 40%, 50%, 80% and 10%.
- Agent Attraction: 10%, 80%, 30%, 70%, 60%, 40%, 50%, 20% and 90%.

In addition, a randomised order for the position of control videos was generated, between either left (L) or right (R):

- Agent Repulsion: R, R, L, R, L, R, L, L and L.
- Agent Attraction: L, L, R, R, L, L, L, R and R.

The stimuli itself was thirty-six recorded video clips of approximately ten second duration, half showcasing the different intensities with regards to the behavioural feature and half as controls without the behavioural feature present, as noted in Subsection 6.3.2. The different stimuli intensities were created through varying the force weights that were calculated from the crowd data. This was discussed in Subsubsection 6.3.2.2, with Table 6.3 mapping the different intensities to their specific weights. It should be noted that the stimuli for 50% intensity is always based on the values from the crowd data, providing a focal point between reality and perception.

An online platform was used to present the experiment to participants, as outlined in Subsection 6.3.3 and shown in Figure 6.13. Each intensity presented to participants is called a trial and on each trial the participants evaluate the virtual footage. They select a percentage value from 0% to 100%, depending on how realistic they find the crowd behaviour. Higher percentages determine it to be more realistic and lower percentages mean the participants found it unrealistic. A percentage value of 50% or higher is considered a seen-response for the purpose of psychometrics and threshold calculation. In addition, this percentage value is also converted to a perceived realism value for finding the highest response and identifying the optimum configurations in terms of the behavioural feature.

The agent repulsion trials were presented first and the agent attraction trials followed. On each trial a participant was shown both the simulated footage with the behavioural feature present and the simulated footage without the behavioural feature present, side by side. They were asked which of the two videos shows more realistic crowd behaviour and had to select a percentage using a slider with a displaying value, in order to submit and move on. Note, as a forced choice experiment, participants had to provide responses for all trials within the duration of the stimuli.

6.3.5 Results

As a 2AFC method was used for this experiment, the absolute thresholds are taken at the 75% seen level. This is due to the nature of the forced choice paradigm and its effect on participants responses. A seen result is classified as a perceived realism response of 50% or higher and a percentage for the total number of seen results was calculated based on the pool of participants. Percentage seen is utilised for psychometrics and in conjunction with the perceived realism values, the highest average responses that fall within the perceptual thresholds can be identified.

With the forced choice approach, participants picked between two videos on each trial. One with the stimuli present and one with a control without the stimuli present. Responses that identified the control as more realistic were discarded for psychometric calculation, however as expected the instances of this occurring were minimal. Based on the pool of participants and the frequency of control selection, a comparison can be drawn based on the percentage of participants who perceived agent behaviour with social forces to be more realistic. 94% of participants found crowd behaviour with the agent repulsion social force present to be more realistic and 95% of participants found crowd behaviour with the agent attraction social force present to be more realistic. Overall, a total of 95% of participants found crowd behaviour with social forces present to be more realistic, when compared to crowd behaviour without social forces present. Given that a social forces model is implemented to improve the realism of crowd behaviour and has been highlighted for various lines of research, this high percentage towards social forces is as expected.

The psychometric function was plotted for both agent attraction and agent repulsion, as seen in Figure 6.14. The graph's percentage seen responses are on the ordinate and stimuli intensity on the abscissa, as is typical in psychophysics. As can be seen from the resolution of the graph with percentage seen beginning at 68% on the ordinate, the responses were high across the entire psychometric function, contributing to the inconsistent shape. With a 2AFC being employed as a classical method rather than using an adaptive method like a staircase procedure, it allowed for the benefit of a wide probing of the psychometric function. Even so, only one absolute thresholds can be seen on the graph at the point where the agent repulsion psychometric function crosses the 75% seen mark on the ordinate for the lower threshold. There are several factors that could cause responses to be high, including the forced choice paradigm with the use of controls. Social forces are a highly visible behavioural feature and in contrast with the videos without social forces, it could potentially drive up responses. In addition, certain graphical elements of the simulation design could have contributed to the narrow range of responses, as agents were represented as basic geometry without specific animations. This could have potentially made the subtle changes in speed and direction more difficult to observe.

As can be seen in Figure 6.14, due to the narrow range of responses, only one absolute threshold could be identified for agent repulsion at 13% intensity. This links to weight value of 0.13, causing approximately 13% of agents to visibly display the force. By keeping at or above this intensity level, the behavioural feature is considered perceptually plausible up to an intensity of 90% or a weight of 0.9. For agent attraction, the absolute threshold can be considered as 10% intensity up to 90% intensity or a weight of 0.1 up to 0.9, based on the whole psychometric function that was probed.

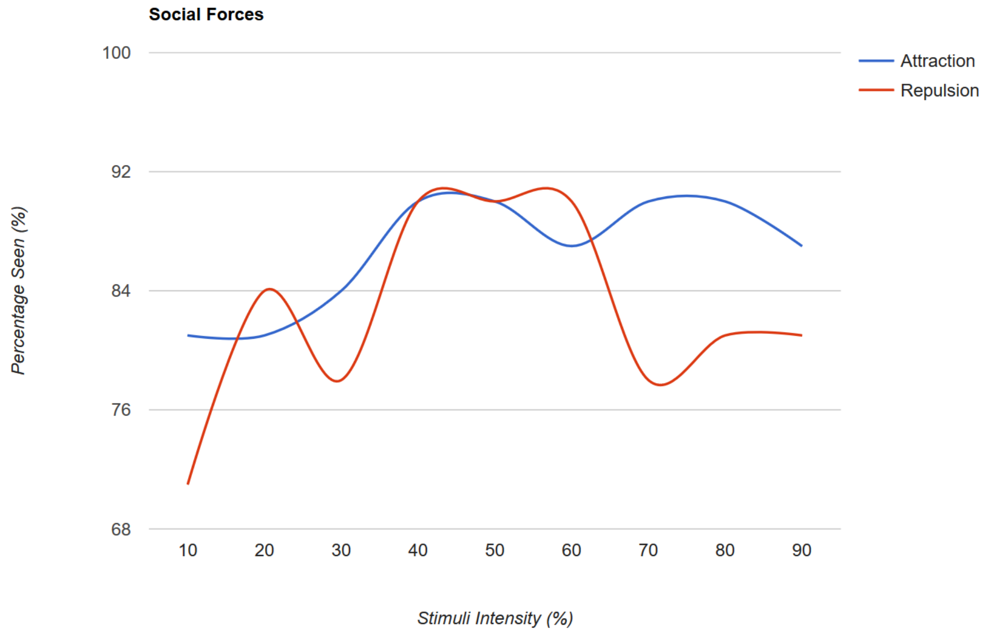

FIGURE 6.14: A graph showing the psychometric functions for the agent attraction and the agent repulsion social forces, with percentage seen responses on the y-axis and stimuli intensity on the x-axis. Note, the axis for percentage seen starts at 68% due to the high responses across the function.

Social Force	Lower Threshold (w)	Upper Threshold (w)	Optimum (w)
Agent Attraction	0.1	0.9	0.7
Agent Repulsion	0.13	0.9	0.4

TABLE 6.4: A summary of the absolute thresholds and optimum configuration for the social forces behavioural feature.

The optimum configuration was identified using the perceived realism values collected through the psychometric data, as can be seen in Figure 6.15. For the agent attraction force, the highest perceived realism value of 0.88 can be seen at 70% intensity, equating to a weight of 0.7. For the agent repulsion force, the highest perceived realism value of 0.86 can be seen at 40% intensity, equating to a weight of 0.4. A summary of absolute thresholds and optimum configuration are shown in Table 6.4. Further discussion and the difference between virtual and perceived realism are presented in Section 6.5.

6.4 Experiment 3: Grouping Dynamics

The third and final experiment presented in this thesis examines the grouping dynamics behavioural feature. The following Subsections will outline the behavioural feature and

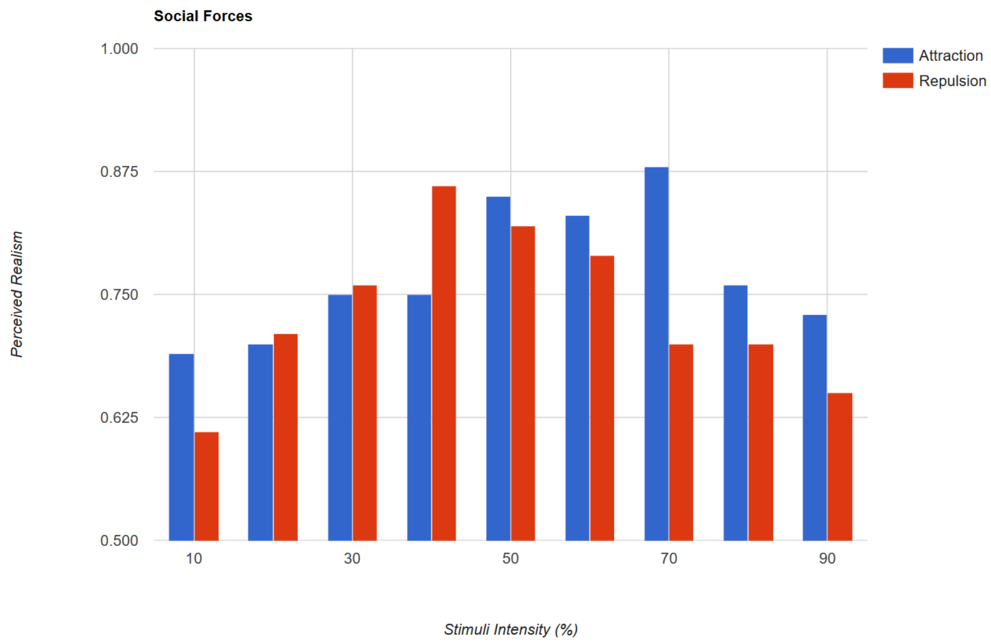

FIGURE 6.15: A graph showing the perceived realism values for both the agent attraction and agent repulsion social forces. Note, the axis for perceived realism starts at 0.5, due to the high responses across the stimuli intensity range.

its variables, the testing apparatus and the number of participants. The experimental procedure is covered in detail and the results are discussed.

6.4.1 Grouping Dynamics

The behavioural feature for grouping dynamics was identified for the final experiment as part of the analysis stage, as discussed in Section 4.1.1. The configuration of a crowd in terms of its groups, is an important aspect for how it behaves. Literature has examined this factor in terms of crowd behaviour in the past(Peters & Ennis 2009), highlighting its significance in terms of simulation. In addition, it is a highly visible feature of crowd behaviour and is applicable to the majority of virtual crowd implementations. This shows why it was easily seen within the real-world video footage and thus identified for this experiment. Due to grouping dynamics being a major consideration for most crowd simulations, it also shows the wide applicability of the resulting perceptual metrics and adds credence to the features selection.

When looking at the individual groups that formulate a crowd, two main variables can be seen:

- **Group Frequency**: The number of groups present within the crowd at a given time step.
- **Group Density**: The number of individual agents present within each group or the overall group cardinality.

By altering these two variables with respect to the groups that make up a virtual crowd, it changes the overall crowd behaviour and affects its perceptual plausibility. By considering this grouping dynamics behavioural feature for psychophysical evaluation, the perceived realism of this important factor is explored. The results yield perceptual metrics in terms of thresholds and optimum configurations, which will provide aid for related research and future implementations of virtual crowds.

6.4.1.1 Psychophysical Method

The psychophysical method selected to assess the grouping dynamics behavioural features was the method of constant stimuli, which falls into the category of a classical method. An additional comparative element was applied to further enhance the base method. Specifics regarding the procedure of applying this method as part of this experiment are discussed in Section 6.4.4.

The main reasons for choosing the method of constant stimuli for this third experiment mirrors many of those as were stated in Section 6.3.1.1 for the social forces experiment. Once again, with further refinement to the urban crowd simulation and the study being of a suitable scale distributed via the online platform, the benefits of the classical method to provide probing of the whole psychometric function or stimuli range, is worthwhile. The element of randomness inherent in the method of constant stimuli is invaluable, especially considering the multiple locations and large number of individual trials. Of the classical methods, constant stimuli has been found to be effective (Simpson 1988) and allows for the psychometric function to be plotted and the absolute thresholds identified.

The main difference between this and the second experiment, apart from the behavioural feature and different locations being tested, is the use of a comparative method instead of 2AFC. This change was influenced in part by the high responses acquired from participants during the social forces experiment. As was outlined, forced choice is known to produce high seen responses and has been applied to areas such as the discrimination of motion perception (Gold & Shadlen 2000) with success. However, in the case of social forces this tendency proved problematic for identifying all of the absolute thresholds and thus was a consideration for this experiment. While forced choice can be a powerful tool when considering multiple stimuli, for the first assessment of a new behavioural feature it proved unsatisfactory and must be carefully considered for future application. The comparative method borrows an element of forced choice in terms of stimuli comparison, however in this instance it feeds back to the general methodology outlined in Section 4.1, to provide a comparison to the original crowd video footage, thus grounding participant responses. This type of comparison between has shown success within similar applications that compare reality and generated content (Peters et al. 2008).

6.4.2 Stimuli Creation

A number of short video clips were required as stimuli for psychophysical evaluation of the grouping dynamics behavioural feature. An example of the final stimuli can be seen in Figure 6.16.

The implementation details regarding this behavioural feature were outlined in Section 5.9. To summarise, the grouping dynamics variables are setup through a group control system in the Unity development environment, which allows for groups to be created and then set with a density. Once the correct number of groups with correct density are ready, the simulation is initialised. Once runtime begins, video capture software records a video for approximately ten seconds. The process is then repeated with a different configuration with respect to the behavioural feature. For this experiment a total of thirty-six videos were recorded at 60 frames per second for stimuli, each with a different intensity of the features variables. In addition, three extra videos, each with a duration of ten second showing the real-world crowd footage, are produced from the pool of crowd data. Each of these three real-world videos are from one locations the virtual scenes are based on. Aside from the two variables being altered, all other parameters and algorithms were kept constant during stimuli creation. Additionally, all groups were set to follow specific paths, meaning positions and the overall direction of individual groups remains constant as well. This ensures that the only visible differences in terms of the crowd behaviour are due the alteration in the intensity of the behavioural feature.

FIGURE 6.16: Examples of the video stimuli recorded for the grouping dynamics experiment, showing Bourbon Street, Temple Bar and 41st Street respectively.

6.4.2.1 Crowd Data

The process of stimuli creation was grounded in reality by analysing crowd data for initial values. This helped to ensure any results were useful for a wide range of different crowd simulation purposes The experiment consisted of the recreation of three real-world locations, which were selected for their crowd behaviour and implementation potential. These three locations, as shown in Figure 6.17, were:

1. Bourbon Street in New Orleans.
2. 41st Street in New York.
3. Temple Bar in Dublin.

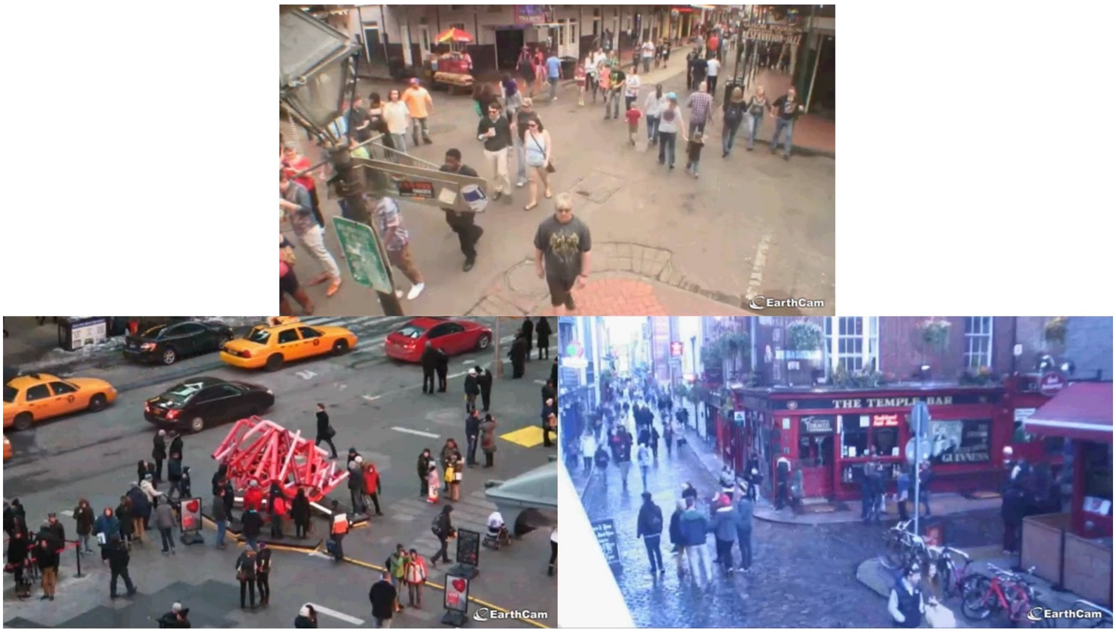

FIGURE 6.17: These three locations provided data with regards to the grouping dynamics behavioural feature and were then recreated as virtual environments for psychophysical evaluation.

With the grouping dynamics behavioural feature as the focus for analysis, a manual annotation was tailored for pedestrian group configurations. Through visual identification at key frame intervals, groups were annotated with their configuration recorded. Figure 6.18 shows an example of this process. The behavioural feature consists of two key variables, group frequency and group density. By annotating the pedestrian groups with respect to these two variables, values can be calculated at each key frame. These values can then be combined and averaged to give overall values for that specific location. Table 6.5 shows these resulting frequency and density values.

FIGURE 6.18: A key frame from one of the selected video clips. This shows an example of the annotation process whereby groups are tagged. The colour of the annotation denotes density of the group.

Location	Frequency	Density
Bourbon Street	20	2.4
41st Street	22	2.3
Temple Bar	16	2.3

TABLE 6.5: The resulting group frequency and group density values as calculated from the real-world crowd footage.

6.4.2.2 Variation

In order to conduct a psychophysical experiment, a number of different intensities of stimuli are required. By analysing crowd data to find initial values for the grouping dynamics variables, these values could then be set as 50% intensity. Variation could then be calculated through the percentages, to provide a wide range of stimuli for the psychometric function. Table 6.6 highlights the results of this process and shows the different intensities and what they relate to in terms of group frequency and group density for each location. Note that values are rounded appropriately, with frequency requiring an integer and density a float to one decimal place. These values link directly to the number of groups and agents in each group. For example, a group frequency of four and a group density of two and a half, would translate to four groups being displayed, two consisting of two agents and two consisting of three agents.

Intensity (%)	Bourbon St. (f)	Bourbon St. (d)	41st St. (f)	41st St. (d)	Temple Bar (f)	Temple Bar (d)
10	4	1.0	4	1.0	3	1.0
30	12	1.4	13	1.4	10	1.4
50	20	2.4	22	2.3	16	2.3
70	28	3.4	31	3.2	22	3.2
90	36	4.3	40	4.1	29	4.1
100	40	4.8	44	4.6	32	4.6

Table 6.6: The resulting intensities linked to the specific group frequency and group density values.

6.4.3 Apparatus and Participants

A total of seventy-eight participants (69 males and 9 females, aged between 18 and 50) from various education backgrounds but all with basic computing skills, took part in this third experiment. All participants were naïve as to the overall purposes of the experiment, but were informed as to what they were judging on each trial.

In order to reach a larger number of participants, a specially developed psychophysics platform was deployed online to improve accessibility. Figure 6.19 shows four webpages from this platform. Upon navigating to the webpage and beginning the test, a participant is presented with instructions and must agree to be part of the experiment to continue. Next some basic demographic data is taken and the experiment begins. A series of trials are then presented to each participant.

6.4.4 Procedure

This experiment examines the grouping dynamics behavioural feature and its two variables, group frequency and group density. The psychophysical method employed was the constant stimuli approach, with a comparative element. Constant stimuli was the appropriate method, as it allows for a wide probing of the psychometric function, while controlling the potential of participants identifying the intensity scale through randomisation of stimuli order. The comparative element, whereby real-world footage is shown alongside its virtual counterpart, aided in focusing participants and providing a reference point to ground responses in reality.

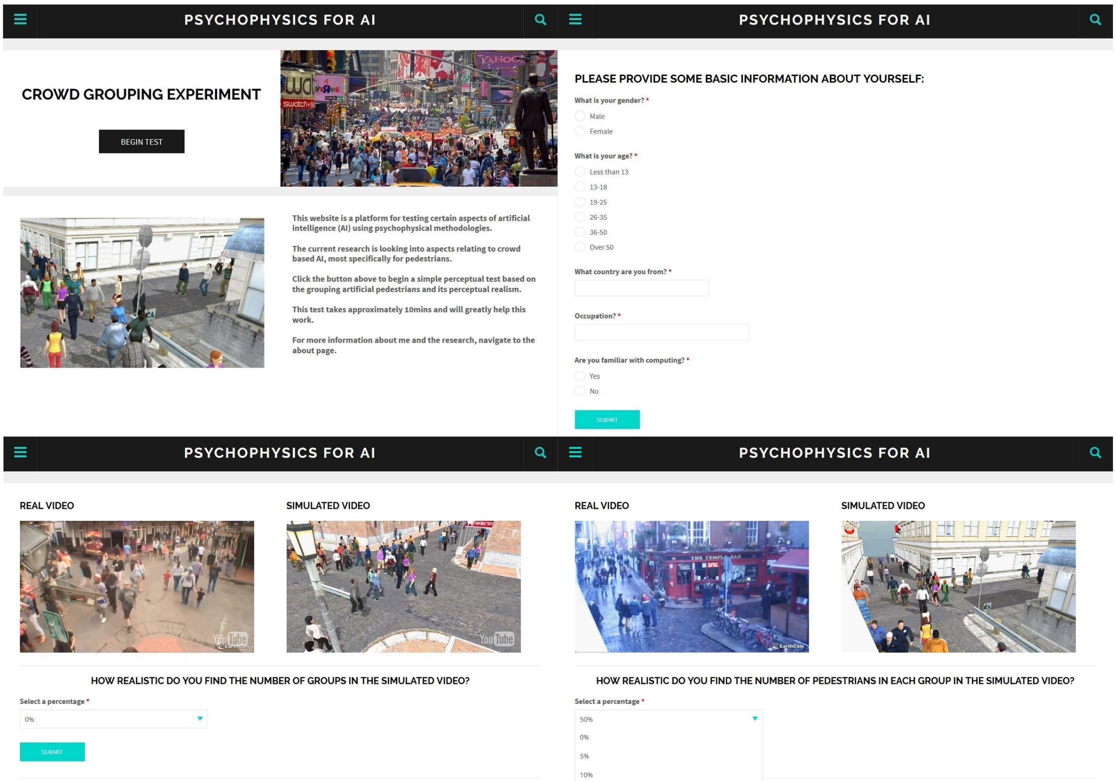

FIGURE 6.19: Four webpages from the online platform. Top left is the home screen, top right collects demographic data, bottom left is a Bourbon frequency trial and bottom right is a Dublin density trial.

Three locations were selected to provide a range of data and examine possibilities that location has an impact on results. Each location was analysed to provide a starting point for stimuli, as outlined in Subsecton 6.4.2.1. Over the three locations, for the two variables, a total of six percentage-based intensities were selected. This allowed for acceptable probing of psychometric function, while also keeping the experiments length within reason so as to not dissuaded participants. These intensities were:

- 10%, 30%, 50%, 70%, 90% and 100%.

By using the method of constant stimuli, a randomised order for presenting the different intensities was generated:

- Location One: 70%, 90%, 30%, 50%, 100% and 10%.
- Location Two: 10%, 100%, 30%, 70%, 50% and 90%.
- Location Three: 90%, 70%, 100%, 10%, 30% and 50%.

The stimuli itself was thirty-six recorded video clips of approximately ten second duration, showcasing the different intensities with regards to the behavioural feature, as noted in Subsection 6.4.2. The different stimuli intensities were created through varying the values analysed from the crowd data, hence different locations and variable intensities map to slightly different values. This was discussed in Subsubsection 6.4.2.2, with Table 6.6 mapping the different intensities to their specific values. It should be noted that the stimuli for 50% intensity is always based on the values from the crowd data, providing a focal point between reality and perception. Three additional video clips were used in the experiment, each showing ten seconds of footage from one of the real-world locations that were analysed.

An online platform was used to present the experiment to participants, as outlined in Subsection 6.4.3 and shown in Figure 6.19. Each intensity presented to participants is called a trial and on each trial the participants evaluate the virtual footage. They select a percentage value from 0% to 100%, in increments of 5%, depending on how realistic they find the grouping. Higher percentages determine it to be more realistic and lower percentages mean the participants found it unrealistic. A percentage value of 50% or higher is considered a seen-response for the purpose of psychometrics and threshold calculation. In addition, this percentage value is also converted to a perceived realism value for finding the highest response and identifying the optimum configurations in terms of the behavioural feature.

The group frequency trials were presented first, showing one location, then the other and so on. Finally, the group density trials were then presented in the same manner. On each trial a participant was shown a real-world video clip on the left and a simulated video clip on the right. They were asked about the realism of the groups in the simulated video and had to select a percentage from a drop down box, before they could submit and proceed to the next trial. Note, the participants were informed as to the type of trial, whether it be group frequency or group density.

6.4.5 Results

The purpose of these results was to highlight both the absolute thresholds and the optimum configurations with respect to perceptual plausibility. In addition, the effect of location with respect to the psychophysical method is considered. For the method of constant stimuli without any forced choice element, the absolute thresholds are taken at the 50% seen level. As with the other experiments, a seen result is classified as a perceived realism response of 50% or higher. The number of seen results were collected and a percentage for the total number of seen results was calculated based on the pool of participants.

This experiment consisted of three locations and as such psychometric functions were plotted for each. These psychometric functions are for the variables group frequency and group density, which form the behavioural feature of grouping dynamics. The psychometric functions can be seen in Figures 6.20, 6.21 and 6.22. For all three graphs, percentage seen responses are on the ordinate and stimuli intensity on the abscissa, as is typical in psychophysics. The overall shape of these psychometric functions were as expected, with a peak in terms of the percentage seen responses that drops off as the stimuli intensity is either increased or decreased away from this optimum level. In addition, all three graphs show that the range of stimuli intensities were broad enough to have allowed for the absolute thresholds to be identified, due to the fact that the point at which the psychometric function crosses the 50% seen mark on the ordinate is visible for all functions for both the upper and lower thresholds.

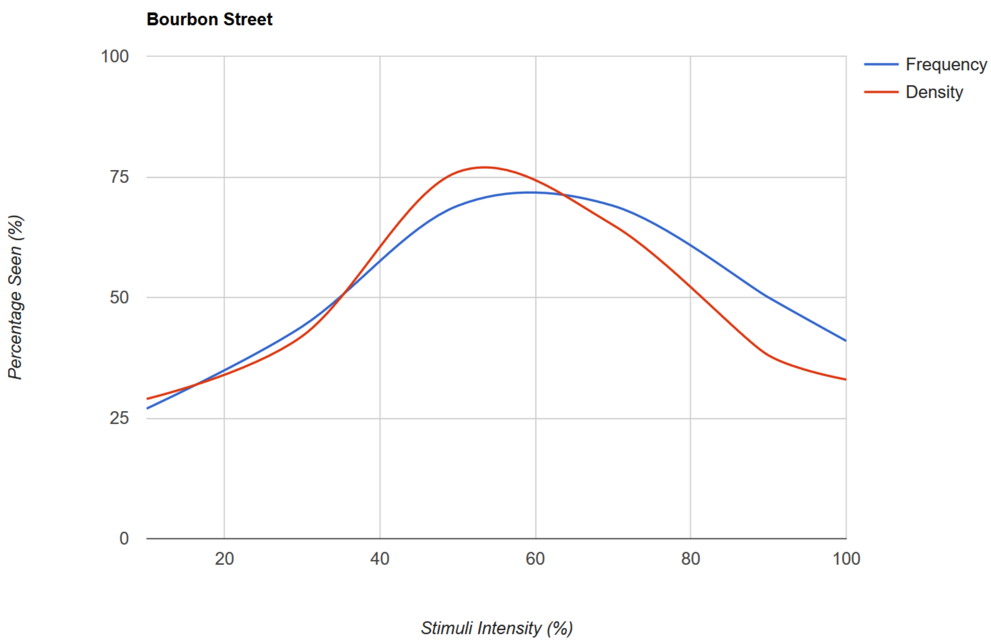

FIGURE 6.20: A graph showing the psychometric functions for group frequency and group density at the Bourbon Street location, with percentage seen responses on the y-axis and stimuli intensity on the x-axis.

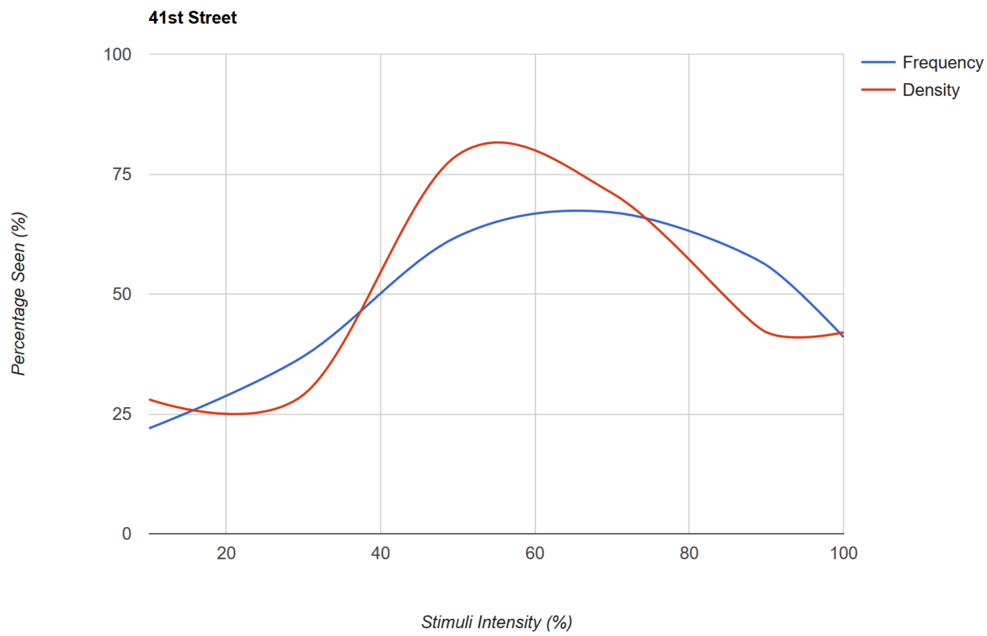

FIGURE 6.21: A graph showing the psychometric functions for group frequency and group density at the 41st Street location, with percentage seen responses on the y-axis and stimuli intensity on the x-axis.

As can been seen in 6.20, the absolute thresholds for group frequency in Bourbon Street were identified as 35% intensity and 90% intensity, which links to a frequency of 14 and 36. In addition, as the graph shows, the absolute thresholds for group density in Bourbon Street were identified as 35% intensity and 81% intensity, which links to a density of 1.6 and 3.9. As can been seen in 6.21, the absolute thresholds for group frequency in 41st Street were identified as 41% intensity and 94% intensity, which links to a frequency of 18 and 42. In addition, as the graph shows, the absolute thresholds for group density in 41st Street were identified as 39% intensity and 84% intensity, which links to a density of 1.8 and 3.8. As can been seen in 6.22, the absolute thresholds for group frequency in Temple Bar were identified as 35% intensity and 95% intensity, which links to a frequency of 12 and 31. In addition, as the graph shows, the absolute thresholds for group density in Temple Bar were identified as 33% intensity and 81% intensity, which links to a density of 1.5 and 3.7.

The optimum configuration was identified using the perceived realism values collected through the psychometric data. As can be seen in Figure 6.23, for group frequency in Bourbon Street, the highest perceived realism value of 0.7 can be seen at 50% intensity, equating to a value of 20 groups. The graph also shows that for group density in Bourbon Street, the highest perceived realism value of 0.81 can be seen at 50% intensity, equating

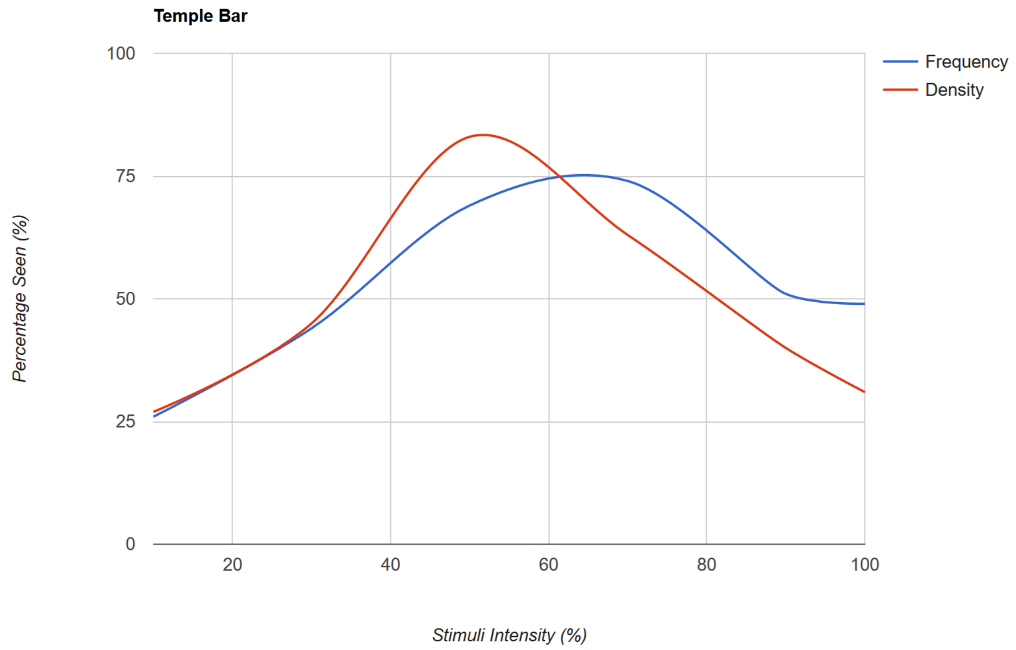

FIGURE 6.22: A graph showing the psychometric functions for group frequency and group density at the Temple Bar location, with percentage seen responses on the y-axis and stimuli intensity on the x-axis.

to a value of 2.4 agent per group. As can be seen in Figure 6.24, for group frequency in 41st Street, the highest perceived realism value of 0.68 can be seen at 70% intensity, equating to a value of 31 groups. The graph also shows that for group density in 41st Street, the highest perceived realism value of 0.83 can be seen at 50% intensity, equating to a value of 2.3 agents per group. As can be seen in Figure 6.25, for group frequency in Temple Bar, the highest perceived realism value of 0.71 can be seen at 70% intensity, equating to a value of 22 groups. The graph also shows that for group density in Temple Bar, the highest perceived realism value of 0.88 can be seen at 50% intensity, equating to a value of 2.3 agents per group.

Two one-way repeated measured analysis of variance (ANOVA) were conducted on the psychometric data to evaluate the null hypothesis that there is no change in participants responses when measured at the three different locations of 41st Street, Bourbon Street and Temple Bar (N=6), for both variables group frequency and group density. The results of the group frequency ANOVA indicated no significant effect, Wilks Lambda = .456, F(2, 4) = 2.386, P ¡ .208, = .544. The results of the group density ANOVA similarly indicated no significant effect, Wilks Lambda = .849, F(2, 4) = .356, P ¡ .721, = .151. Thus, there is no significant evidence to reject the null hypothesis.

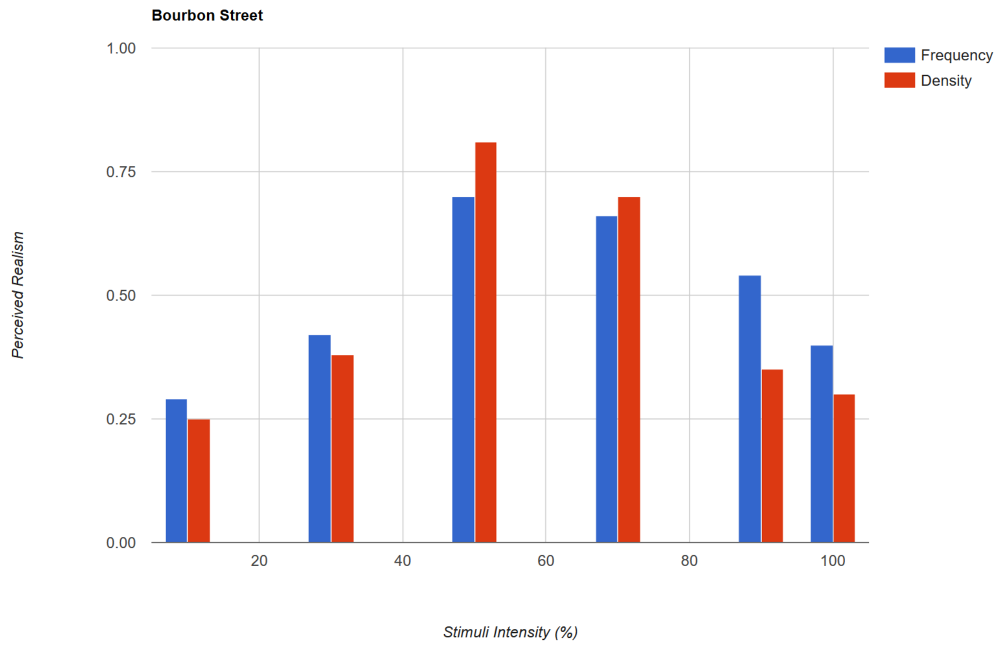

FIGURE 6.23: A graph showing the perceived realism values for both group frequency and group density at the Bourbon Street location.

Grouping Dynamics	Lower Threshold	Upper Threshold	Optimum
Frequency (Bourbon)	14 Groups	36 Groups	20 Groups
Density (Bourbon)	1.6 Per Group	3.9 Per Group	2.4 Per Group
Frequency (41st)	18 Groups	42 Groups	31 Groups
Density (41st)	1.8 Per Group	3.8 Per Group	2.3 Per Group
Frequency (Temple)	12 Groups	31 Groups	22 Groups
Density (Temple)	1.5 Per Group	3.7 Per Group	2.3 Per Group
Frequency (Average)	15 Groups	36 Groups	24 Groups
Density (Average)	1.6 Per Group	3.8 Per Group	2.3 Per Group

TABLE 6.7: A summary of the absolute thresholds and optimum configuration for the grouping dynamics behavioural feature.

The average absolute thresholds for group frequency were calculated at 15 groups and 36 groups. The average absolute thresholds for group density were calculated at 1.6 agents per group and 3.8 agents per group. The average optimum configuration for group frequency was calculated at 24 groups and the average optimum configuration for group density was calculated at 2.3 agents per group. A summary of absolute thresholds and optimum configuration are shown in Table 6.7. Further discussion and the difference between virtual and perceived realism are presented in Section 6.5.

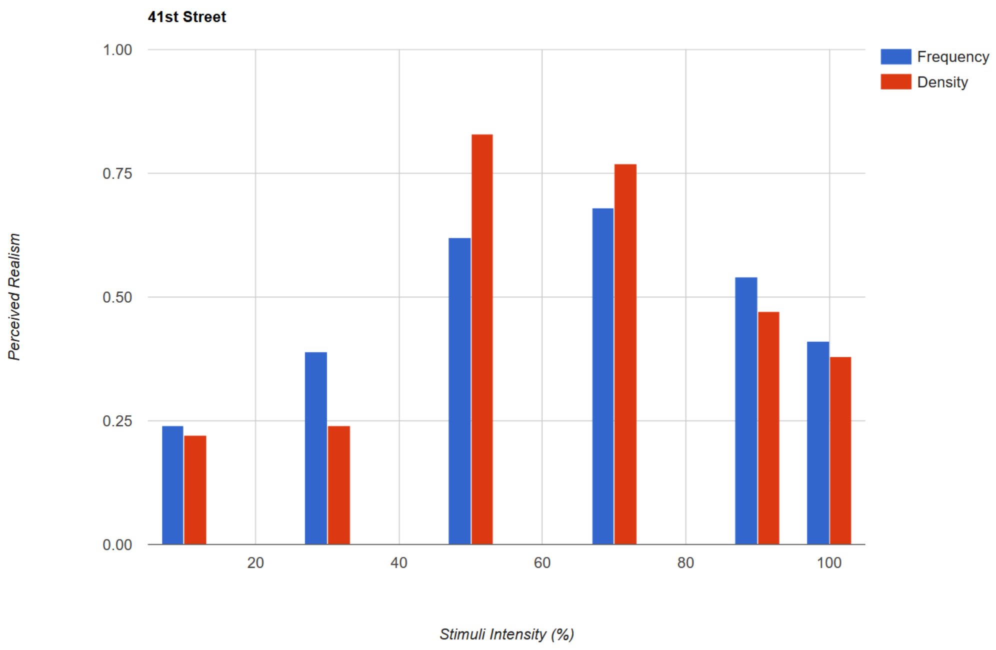

FIGURE 6.24: A graph showing the perceived realism values for both group frequency and group density at the 41st Street location.

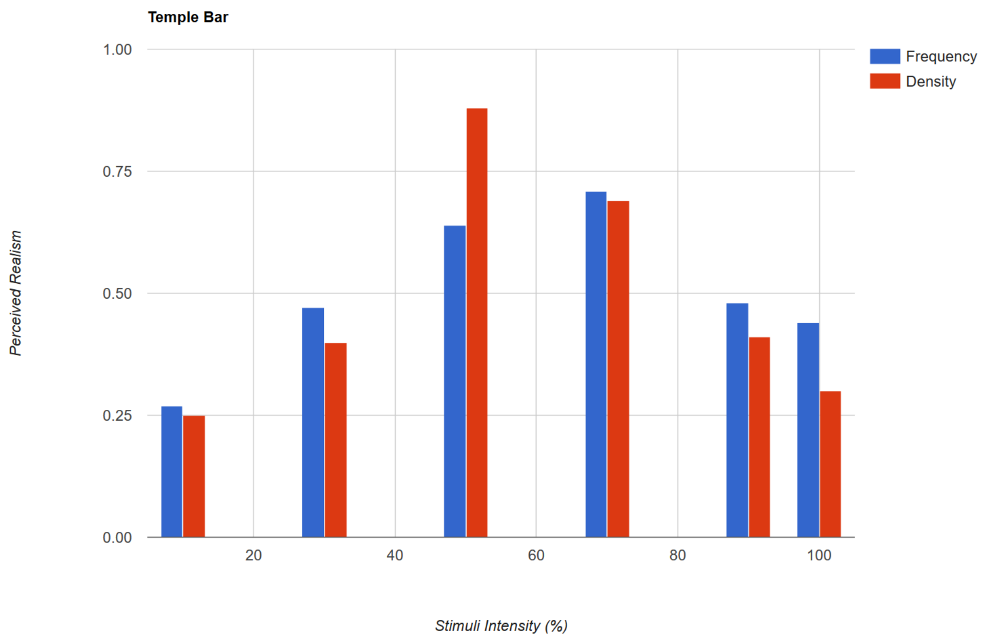

FIGURE 6.25: A graph showing the perceived realism values for both group frequency and group density at the Temple Bar location.

6.5 Discussion

By using psychophysical methods to assess behavioural features, it has been possible to identify the absolute thresholds and optimum configurations for a number of elements

Behavioural Elements	Lower Threshold	Upper Threshold	Optimum
Velocity Range	0.4875 m/s	1.2625 m/s	0.75 m/s
Velocity Distribution	0.725 m/s	1.75 m/s	1.25 m/s
Agent Attraction Force	0.1 w	0.9 w	0.7 w
Agent Repulsion Force	0.13 w	0.9 w	0.4 w
Group Frequency	15 Groups	36 Groups	24 Groups
Group Density	1.6 Per Group	3.8 Per Group	2.3 Per Group

TABLE 6.8: A summary of the absolute thresholds and optimum configuration for the three behavioural features assessed in this research.

important in crowd simulation development. Table 6.8 shows a summary of these perceptual metrics, which are applicable to a wide range of future crowd implementations and research, due to the fundamental nature of the behavioural features. This was a key purpose of employing the three-stage general methodology of analysis, synthesis and perception, in order to identify the widest spread features as a priority. For crowd simulation development, knowing the optimum parameters for achieving perceptual plausibility with respect crowd behaviour can assist in a simulation achieving its purpose. The absolute thresholds also allow developers to scale back certain parameters, while still being in a position that achieves suitable perceived realism. An example of how these thresholds would be useful is for mobile development, where resources are limited. By knowing that 15 groups will not be optimum but will still to achieve a reasonable level perceptual plausibility, a developer can scale back the number of agents on display at once and reduce the amount of computational power required.

In addition to these perceptual metrics however, a key element of this research is to highlight similarities and differences between perception and reality. By using crowd data for each experiment to inform the initial stimuli intensities at 50%, it provided not only a starting point for variation, but a link back to virtual realism. By looking at the responses from participants in the form of the absolute thresholds and optimum configurations, but with respect to the stimuli intensities these represent, it can be seen how perceived realism compares to virtual realism at 50% intensity. Table 6.9 highlights the results in terms of the intensities they represent.

The first major observation is that 50% intensity falls within the absolute thresholds for all of the experiments. This shows for these experiments that virtual realism was always within the realms of acceptable perceived realism. This is an important consideration, as if this was found to not be the case, then it would suggest a disparity between

Behavioural Elements	Lower Threshold	Upper Threshold	Optimum
Velocity Range	19% Intensity	51% Intensity	30% Intensity
Velocity Distribution	29% Intensity	70% Intensity	50% Intensity
Agent Attraction Force	10% Intensity	90% Intensity	70% Intensity
Agent Repulsion Force	13% Intensity	90% Intensity	40% Intensity
Group Frequency (Bourbon)	35% Intensity	90% Intensity	50% Intensity
Group Density (Bourbon)	35% Intensity	81% Intensity	50% Intensity
Group Frequency (41st)	41% Intensity	94% Intensity	70% Intensity
Group Density (41st)	39% Intensity	84% Intensity	50% Intensity
Group Frequency (Temple)	35% Intensity	95% Intensity	70% Intensity
Group Density (Temple)	33% Intensity	81% Intensity	50% Intensity

Table 6.9: A summary of the absolute thresholds and optimum configuration with respect to stimuli intensity.

what is actually realistic and what the viewer considers realistic. This would then cause developers to have to choose between crowd behaviour being either perceptually plausible or virtually realistic. As the results show that virtual realism has fallen within the perceptual thresholds for these behavioural features, a simulation can achieve both types of realism, albeit often not in an optimum manner. This pattern and the nuances of human perception suggest that other behavioural features will similarly show a link between the virtual and perceived realism. As some crowd simulations can require both types, such as for educational purposes, it is an avenue for further exploration.

For the varying velocity behavioural feature, it can be seen that the optimum configuration for velocity distribution correlates with the 50% intensity stimuli based on the crowd data. This shows that in this instance perceived realism links with the virtual realism. For velocity range however, the optimum configuration was identified at 30% intensity, which suggests that viewers favour a narrower velocity range compared to what was present in the crowd data. To further support this, the upper absolute threshold was only 1% above the initial intensity, whereas the lower threshold showed a 31% intensity drop.

For the social forces behavioural feature, the optimum configurations for both agent attraction and agent repulsion present different intensities than 50%. The optimum configuration for the agent attraction force was identified as 70%. By increasing the intensity of this force it causes more agents to be move together and form groups. This suggests that viewers favoured more grouping in terms of the social forces. The optimum configuration for the agent repulsion force was identified as 40%, which through decreasing the intensity would cause less agents to actively move away from other agents. This again suggests viewers favoured a higher percentage of agents grouping together. The intensity ranges given by the absolute thresholds highlight the flexibility of social forces in terms of perceptual pliability, in contrast to the virtual realism.

For the grouping dynamics behavioural feature, the majority of the optimum configurations for both velocity range and velocity distribution at the three different locations, correlated with the initial 50% intensity based on reality. The only two that varied were the intensities for 41st Street group frequency and Temple Bar group frequency, with viewers favouring an increased intensity of 70% for both. In addition, the upper absolute thresholds possessed a bigger difference from 50% intensity, when compared with the lower thresholds. This suggests, along with the two optimum configurations at 70%, viewers had more perceptual flexibility in terms of intensity when it is increased from the initial 50% rather than decreased. It can also be seen that all of the group frequency upper thresholds are higher than their density counterparts, suggesting more viewer tolerance for additional groups rather than increased group size.

Overall these observations suggest a link between virtual realism and perceived realism in terms of crowd behaviour. For all of the optimum configuration intensities, none is more than 20% removed from the initial 50% intensity derived from the real-world crowd data. In addition, it can be seen that all of the 50% intensities fall within the perceptually identified absolute thresholds. While it may be considered that perceived realism is more flexible than virtual realism, both are important considerations and play an important role in crowd simulation. Through future experimentation, these links could be further cemented and applied to various aspects of development. This research and these experiments have presented a new successful view, considering both perception and reality for evaluating features of simulation.

6.6 Summary of Chapter 6

In this chapter, the three experiment conducted for this thesis have been presented. Details have been discussed with regards to each behavioural feature, the various intensities of stimuli produced and how they were applied as part of the experimental procedure. The results have been outlined and discussed, highlighting links between reality in the form of crowd data and human perception when viewing the simulated crowd behaviour. Perceptual thresholds have been calculated and optimum configurations identified, to act as guidelines for future related research and to aid in the development process of virtual crowds. In addition, the experiments have been discussed as a whole, with the similarities and differences between each noted.

In the next chapter, overall conclusions and avenues for future work are discussed. A summary of the contributions made in this thesis is presented, highlighting four areas. Namely, experimental results, perceptual metrics, adapted methods and crowd platform. A reflection is given for the various methods employed in this thesis. Finally, potential areas for future research are considered, looking at ways to conduct more iterations of the general methodology and to improve aspects highlighted in this research.

Chapter 7

Conclusions and Future Work

In this chapter, a summary of the contributions made in this thesis are outlined, along with a discussion of the limitations of the research. A reflection on the general methodology and psychophysical methods employed is provided. Finally, avenues of potential future work are considered.

7.1 Summary of Contribution

The key contribution from the work presented in this thesis is a set of perceptual metrics for evaluating and optimising crowd simulation as part of a proof-of-concept application of a set of methods that can be used to determine appropriate metrics in other scenarios. The work is supported by experimental results and built upon a new crowd simulation platform created for this research.

For the varying velocity behavioural feature, the optimum velocity range and distribution were identified for achieving the highest rate of perceptual plausibility with regards to a crowd's velocity magnitude. In addition to the optimum values, the perceptual thresholds were also identified, showing the limits at which these variables can be altered while still achieving perceptual plausibility. Velocity being the typical metric to provide agents with motion, allows these perceptual metrics to be influential in crowd simulation development. Depending upon simulation requirements, the optimum values and thresholds can be used as guidelines for ensuring perceived realism, in terms of the crowd behaviour for varying velocity. These perceptual metrics will also be influential in future psychophysical experiments applied to crowd behaviour, through informing construction and as part of a corpus of perceptual data.

For the social forces behavioural feature, the optimum weight values were identified for the two major social forces between agents, namely attraction and repulsion. In addition, the perceptual thresholds for these two forces were examined. Despite some slight decreases at higher and lower intensities only one threshold was found, showing social forces can be altered by a large degree while still retaining perceptual plausibility. With social forces being prominent in the literature as an algorithm for incorporating realistic crowd behaviour, this result was not unexpected and shows the applicability of these perceptual metrics for future crowd simulation implementations using a social forces model.

For the grouping dynamics behavioural feature, the optimum group frequency and group density for virtual crowds were identified for three different locations. Additionally, the perceptual thresholds were calculated showing the limits that both group frequency and group density can be altered by, while still achieving perceptual plausibility. Crowd grouping is a highly important element and present in virtually all crowd simulation models. This means that these perceptual metrics can be applied to inform future crowd simulation developments, by showing what is needed in terms of grouping behaviours. Research focusing on agent grouping has been shown in the literature, meaning that these perceptual metrics will be useful considerations in future experiments employed on this feature.

The general methodology applied in this research, consisting of analysis, synthesis and perception, can be applied to a wide range of future perceptual studies investigating elements of simulation or specific algorithms. The process of analysing real-world data to ground a study in reality is of benefit to a wide range of perceptual and psychophysical experiments. In addition, by using a minimalised simulation platform to synthesise stimuli for a specific feature or algorithm, it is possible to focus the perceptual evaluation on that element while reducing potential influence from other factors. Finally, through perceptual evaluation, insights not necessarily considered can be gained on algorithms and features otherwise unevaluated in terms of human perception. The fact that this methodology is applied in an iterative manner, allows for a process of gradual refinement often not seen in a field of bespoke simulations and for a corpus of perceptual data to be built-up that systematically influences design choices. This thesis outlines, defines and then applies this adapted methodology towards crowd simulation and in doing so has shown its advantages. Future research can utilise the methodology for new lines of research, whether towards virtual crowds or another aspect of simulation.

This research has adapted certain psychophysical methods for the purpose of gauging human perception in relation to perceived realism. Psychophysics are classically applied to physical stimuli, such as weight or light perception. By adapting psychophysics to consider perceptual realism, the experiments have been applied to stimuli varied in a virtual sense. By showing through the perceptual metrics discovered in this research that psychophysics can be applied to specialised components such as crowd behaviour through feature identification, different scenarios and future lines of research can apply these methods to other areas of simulation. As noted in the analysis of the literature, the evaluation of algorithms and features of simulation are not prominent compared to other lines of research. By highlighting psychophysics as a possible method for evaluation, it is hoped these lines of research will gain more saturation within the field.

While not the focus in terms of contribution, the experimental results from the three psychophysical experiments presented in this thesis have provided new insights into the links between reality and human perception. These results are important as they not only help to improve knowledge for perceptual plausibility in general, but serve as consideration for virtual crowd behaviour and other similar lines of research. The results from psychophysically evaluating the varying velocity behavioural feature, provided insights into human perception towards the magnitude of velocity ranges displayed by virtual crowds. It was highlighted that viewers favour a narrow range for velocity magnitudes. In addition, the distribution of velocity magnitude favoured the value seen in the real-world crowd data, showing a link between viewer perception and reality. The results from psychophysically evaluating the social forces behavioural feature, provided insights into human perception towards the perceptual plausibility of crowd behaviour when social forces are taken into consideration. It was seen that by incorporating social forces in addition to the core algorithms, the perceived realism of the scene was improved. This shows that viewer perception was in-line with that of reality, however these perceptions were flexible when it came to the specific influence of each social force based on the perceptual thresholds. Finally, the results from psychophysically evaluating the grouping dynamics behavioural feature, provided insights into human perception towards the frequency and density of groups present within crowds. It was shown that viewer perception was in-line with reality, as they favoured the configurations that were based on the values acquired from the crowd data. However, it was also highlighted that viewers had more flexibility towards group frequency, which could remain perceptual plausible for longer than group density, when the intensity was increased. It was noted that viewer perception was impacted more when fewer agents were displayed, when compared to responses when the number of agents were increased.

Again, while not part of the focus for contribution, the urban crowd simulation developed as part of this research has highlighted a pipeline for crowd simulation, which is extensible through parameterisation and the incorporation of behavioural features. The urban crowd simulation was an integral platform for conducting psychophysical evaluation and a core element in validating the methods proposed in this thesis. It is true that this thesis does not focus on the development of novel algorithms, but instead parameterises current algorithms for an implementation that is customisable for the purposes of developing stimuli. Typically, crowd simulations are bespoke applications, developed and applied for a specific scenario or purpose. By outlining a pipeline consisting of core algorithms that can then be tailored towards specific behaviour, it offers an alternative to developers who are looking for similar degrees of customisation and flexibility with regards to crowd simulation.

7.2 Limitations

The scope of this research is for crowd simulations in which success is measured by perceptual plausibility, perceived realism and human acceptance. In a broad sense, this gives a lot of applications for not only the set of methods themselves, but also the discovered perceptual metrics for crowd simulation within typical urban environments. Applications within the field of video games, serious games, simulations for learning and others that utilise crowd technologies intended for human viewing or interaction. While it is true that the perceptual metrics for the three behavioural features are limited in terms of generalisation and cannot be applied to any crowd simulation or scenario, they will be applicable to a good number due to the selection of the behavioural features and their prominence in typical crowd simulation development, along with the selection of the urban environment as a focal point, which features prominently in the majority of instances for virtual crowd. It can be suggested however, that these types of perceptual metrics and the sets of methods are limited by their design and intention. They are not useful in most cases of serious applications for virtual crowds, such as for urban planning or evacuation procedures, as they are based around assessing and optimising around perceptual subjectivity and not data or simulation accuracy. Overall however, the scope is both niche and broad in many respects, as being able to identify perceptual thresholds allows a feature to maintain perceptual plausibility while being able to be tweaked for factors such as performance gains. Particularly, the set of methods is highly adaptable for various uses and it has potential application outside of just crowd-based simulation.

Further on the limitation of the perceptual metrics for generalisation with other crowd simulations is apparent from cultural implications. When considering human perception, cultural bias is a great concern, not only for this research by many others that aim to create metrics or divine specific solutions to widely applicable problems in different arenas. For example, cultural differences between Western Europe and the Middle East are drastic, so it is likely that perceptual metrics regarding crowds would possess visible differences. Indeed, this backed up by the observation that factors such as typical pedestrian walking speed varies between countries and cultures (Levine & Norenzayan 1999), meaning in addition to culture, the specific country or location can also play a great role on these types of metrics. Now, while it is certainly true that in these types of situations, the metrics identified in this thesis would not be application without a degree of inaccuracy, many considerations were put in to place when choosing the specifics of the behavioural features and the locations of the crowd footage used to support the perceptual experimentation, to ensure that they are applicable to a wide range of potential areas. All the crowd footage utilised was based in either the UK or the US, which are heavily urbanised and pedestrianised countries that are often used as a basis for these types of simulations. The US features as the main locations in a massive number of video games, with other fictional locations often based upon the architecture and people found therein. As such, while the application of the metrics to different locations and cultures is certainly a limitation, this was known and care was taken to ensure the perceptual metrics would be as widely useful as possible. In terms of cultural and locational differences for crowd behaviour and its perceptions, it is certainly a potentially rewarding route for future research and the further application of the set of methods presenting in this thesis, which are suited toward this type of a study.

One final noteworthy potential limitation for the perceptual metrics is the scene context in which the experiments were conducted. This includes aspects such as view point, camera location, pedestrian orientation and so on. While it is possible for these to have some impact on the wider generalisability of the perceptual metrics, it should be noted that again this was taken into account as part of the experimental design. By utilising real-world crowd data and in the case of the grouping dynamics experiment, modelling the virtual scene based on the specific view point in which that data was shown, helps to ensure objectivity in the perspective that is being shown and limit influence. To support this idea, research was conducted to perceptually evaluate 'real' pedestrian orientations within a virtual scene modelled based on real-world crowd data, that was compared to generated context and no-context orientations within the same scene (Peters et al. 2008). The study found that the 'real' orientations were found to be real significantly more time of the time when compared to the generated versions with either context or no-context. As such, by incorporating real-world crowd data for modelling and stimuli generation at even low levels, it has helped to mitigate influence from these types of scene context factors.

In terms of the experiments, participants are always a limiting factor for the majority of research endeavours. This holds true for psychophysics, with the methods being perception based and quantitative, a greater pool of participants is always better however that is often not possible given time and various other constraints. Given that, even though more participants would have potentially led to greater accuracy in terms of the results, the numbers of participants for each experiment was within acceptable levels as comparable with other psychophysical and perceptual based research (Peters & Ennis 2009, McDonnell et al. 2008, Gu et al. 2010). In addition, for this research, a specially developed online platform for running these experiments was utilised as an effort to gain as many participants as possible and in mitigate this limitation. In the end, the efforts proved successful and positive results were gained in line with notable literature.

Data collection for all the experiments was primarily from Western Europe and those between 19 and 35 years of age. This ties back into the cultural aspects previously mentioned and again shows the results would likely be inaccurate being applied to such areas and widely different demographics, however the benefit lies in the fact that the data is highly applicable to key westernised urban locales, such as West Europe and the US, where a lot of crowd simulation is utilised and directed at those who fall within the ages of the participants.

As with most research, time and cost can factors prove some of the biggest limitations for the potential scale of the work. The urban crowd simulation in particular was developed from the ground up as part of this research and as can be seen through the development cycle as outlined in Chapter 5, it went through many revisions in C++ OpenGL with refinement being added over time, until the version used for the third experiment was ported to Unity, graphically overhauled and refined into the final version utilised in this research. In an ideal situation that would be the version used for all the experiments to provide more comparable results, as for the social forces experiment it is possible the simplistic graphics and animations contributed to the high responses from participants. However, as is the case with time and cost factors, this was not possible. Instead the general methodology was refined around the concept of iterations of analysis, synthesis and perception, which led to the systems being influenced through real-world crowd data and steady measured refinement process for the urban crowd simulation. It is in the application of the general methodology utilised within this research that these limitations became in many ways benefits, contributing to the successful results. Section 7.3 provides a reflection on the general methodology and the psychophysical methods utilised in this work.

Again, time and cost factors can prove to be some of the biggest limitations for almost any research. It would have surely enhanced results to provide more location comparisons, especially those within different cultures or consider more psychophysical methods to enhance current results with further experiments. In addition, more behavioural features or algorithms could have been explored to provide a more comprehensive list of perceptual metrics and further test the set of methods and framework presented herein. However, within the scope of this doctoral thesis, these were not possible within the time frame and must be considered as avenues for future work, as discussed in Section 7.4. Overall many of the limitation inherent in this work were foreseen and attempts to mitigate their effects on the results were made, which allowed for a positive outcome and realisation of the research's main objectives.

7.3 Reflection

Over the course of this research, the general methodology has been applied and a variety of different psychophysical methods have been employed for experimentation. Both the general methodology and psychophysical methods have been tailored for the specific of this research and the indivdiual experiments, however some reflection on these aspects will provide some potentially useful insights for relevant future work.

7.3.1 Reflection on the General Methodology

The general methodology was applied throughout this research and served as the backbone of the various studies. Being iterative the overall methodology was conducted a total of three times, following the flow of analysis, synthesis, perception and then repeating. The methodology proved to have advantages when applied for the purposes of perceptual evaluation by using psychophysics. These advantages include grounding the perceptual studies in reality, an important consideration for gaining applicable results and initial values for variation with respects to stimuli. In addition, the iterative process allowed for refinement to the urban crowd simulation to be carried out in a methodical, while preserving the initial platform consisting of the core algorithms. Finally, by identifying behavioural features it was possible achieve the focus required for psychophysical evaluation. Overall the methodology proved robust and has potential to be applied to future research where perception is a core element for evaluation.

7.3.2 Reflection on the Staircase Procedure

A staircase procedure was applied for the varying velocity experiment. Staircase procedures are adaptive methods, which are tailored in terms of how the stimuli is presented by the responses from the participant. While this is controllable in small scale studies, on a larger scale it would involve the generation of far more stimuli than would be required for classical methods such as constant stimuli. In addition, this type of method focuses on the area surrounding a threshold. In a typical psychophysical experiment this can prove useful, as they are detection exercises that produce a sigmoidal curve for the psychometric function; however, in the adapted methods applied in this research and the potential for two thresholds rather than just one, a method that provides a wider view of the psychometric function is better suited. As noted, due to the fact staircase procedures do focus on threshold areas, it would prove useful to apply this method after the full psychometric function has been outlined.

7.3.3 Reflection on the Method of Constant Stimuli

The method of constant stimuli was applied for the social forces experiment and the grouping dynamics experiment. Constant stimuli proved a useful method for these experiments, as it allowed for data to be collected across the psychometric function, rather than entirely focusing on an area around the threshold. These experiments were intended to probe perceptual realism and due to the nature in which the stimuli was varied based on the crowd data, upper and lower thresholds could be seen in relation to the perceptual plausibility. By increasing stimuli intensity from the original 50% mark, it was possible to find a point at which the stimuli stopped being perceptually plausible, as the intensity got to high. This was the same in terms of lowering stimuli intensity. Due to a common psychometric function being sigmoidal in shape for physical stimuli detection, the act of focusing on a threshold can be useful; however, due to the adapted nature of the psychophysical experiments conducted in this thesis, a wider view of the psychometric function preferable. In addition, it allows for better identification of optimum configuration values, which are an important consideration for future implementations of virtual crowds.

7.3.4 Reflection on 2AFC

The 2AFC psychophysical method was applied for the social forces experiment. Forced choice causes a participant to have to select between two stimuli promptly. This method varies from the others in the respect that a control can be incorporated and participants are forced to choose between different stimuli. This method is useful for experiments where more than one stimuli needs to be displayed, or when asking participants to pick between choices. The social forces experiment used a control group, to have participants pick between a simulation with the specific social force present and one without. Overall it proved useful for judging the perceptual value of a behavioural feature with regards to the base simulation and has potential use for future experiments with different conditions, however for the social forces experiment it potentially contributed to the high level of responses and thus the range of stimuli intensities must be careful considered or preselected through previous experimentation.

7.3.5 Reflection on the Comparative Method

The comparative method was applied for the grouping dynamics experiment. This method is not typically utilised in psychophysical experiments, but is a construct of this research. By applying this method, a video clip of the original crowd footage is provided at each trial, side by side with a stimulus. This acted as a reference point for participants. The method proved useful, allowing for the successful identification of both thresholds and optimum configurations. The comparative method has potential for further exploration in future experiments.

7.4 Future Work

The general methodology of analysis, synthesis and perception was applied for this research. It is an iterative methodology, meaning it is intended to be repeated to formulate an ever broader corpus of perceptual results. In this thesis, three iterations of this methodology were completed, leading to the psychophysical evaluation of three behavioural features. The most obvious area for future work is to continue adding to the corpus of perceptual results, through the identification and evaluation of new behavioural features. There are a number of areas for possible future work in this regard.

An area for potential future lines of research is with the behavioural annotation mechanism. This method of implementation has been seen in the literature (Peters & Ennis 2009, Shao & Terzopoulos 2005, Peters et al. 2003, Anderson 2003) and can be utilised for tailoring specific crowd behaviours through environmental features. It was noted that a behavioural annotation system was implemented for the urban crowd simulation, but it was not required for the identified behavioural features presented in this thesis. By analysing real-world crowd behaviour with respect to environmental features such as roads, pedestrian crossing and so on, it would be possible to inform synthesis for behavioural annotations to provide environmentally triggered behaviours. In doing so it would then be possible to employ the 2AFC psychophysical method, to perceptually evaluate if these specific annotated behaviours improved the perceptual plausibility of crowd behaviour within the scene. The 2AFC method specifically, would allow for participants to make a choice between the different annotations and between scenes with annotations present and those without. In this way a new corpus of perceptual data could be created towards specific annotations, providing implementation guidelines with respect to the effectiveness of each annotation, their optimum values and the perceptual thresholds. By further highlighting the potential of using psychophysical methods for evaluating algorithms and specific features of simulation, it is of benefit to the area and serves to improve saturation, while also promoting the use of behavioural annotation as a method for implementing specific crowd behaviours.

Another avenue of future research would be to consider how the physical attributes of agents affect the viewer's perception of the behaviour they expect to see. This type of implementation would not be to dissimilar from conditions considered by the HiDAC system (Pelechano et al. 2007, Pelechano & Badler 2006), which considers agents psychological, physiological and social factors in its model. By analysing the behaviour of pedestrians with certain physical attributes, it would be possible influence behavioural design for different agent variables. For example, a pedestrian over a certain height, could be influenced to walk faster or vice versa. Probing the perception based around this in terms of crowd behaviour would likely produce some interesting results. The psychophysical method that would be most appropriate for this study would be the method of constant stimuli, in order to fully explore the psychometric function. In addition, utilising the comparative method would provide a point of reference for participants to focus responses, as seen with the experiment on grouping dynamics. This study would not only provide guidelines for achieving perpetual plausibility with regards to implementing different physical attributes for agents, but it would also highlight some of the links in perception between physicality and behaviour. These insights could potentially be useful for various lines of research and as considerations for simulation.

These examples of future work for additional iterations of the general methodology show the untapped potential of applying psychophysical methods for evaluating components of simulation. It shows the benefits of conducting such studies and how they contribute to the overall body of research. In addition to these examples, there were some areas highlighted in this research that could be applied for future work.

One area for additional work can potentially be seen in the results of the social forces experiment. In this experiment only one perceptual threshold was identified, showing that an additional investigation could yield yet more results. The method of constant stimuli was used to give an overview of the psychometric function, however a more tailored approach through an adaptive psychophysical method such as a staircase procedure, would give results more focused towards the area surrounding the thresholds. By conducting an additional study using a staircase procedure, the new results could enhance the previous results to provide a more in-depth look at this specific behavioural feature. This could potentially allow for the identification of additional thresholds and supplement the existing work. This additional study could also be applied to the other behavioural features, again to add more depth to the results. While this provides an interesting avenue for possible future work, it was not practical for this thesis as the results were positive, highlighting optimum configurations and thresholds in most cases. The responses for social forces were high and it is possible a further study might provide additional insight, however there is no guarantee.

The process of analysis is currently conducted manually and it is a time consuming process. This shows that an area for potential future research could be to apply a form of automatic analysis with respect to the real-world crowd footage. This would streamline the analysis stage of the general methodology, bringing the added benefit of being able to further supplement the corpus of perceptual data beyond what is currently possible. Current technology exists for video content analysis to detect certain temporal and spatial events. It would no doubt be possible to apply these technologies towards crowd behaviour displayed within real-world video footage. Indeed, research has already considered applying these technologies for the analysis of crowd data (Kok et al. 2016, Dehghan et al. 2014).

One final aspect for consideration in terms of future work is the adaptation of more psychophysical methods. This would allow their effectiveness for gauging perceived realism in terms of crowd behaviour to be judged. The current methods that have been adapted include the more prominent methods of psychophysics, such as the method of constant stimuli, staircase procedures and 2AFC; however, it is also possible that other methods, such as the method of adjustment and magnitude estimation, would useful as well. An adapted version of magnitude estimation in particular would be applicable to experiments looking to rank behavioural features for effectiveness in terms of perceptual plausibility. This is a possible future line of research that could yield interesting results, once the corpus of perceptual data has been built-up to include enough behavioural features to make the exercise worthwhile.

Bibliography

Almeida, J. E., Rosseti, R. J. & Coelho, A. L. (2013), 'Crowd simulation modeling applied to emergency and evacuation simulations using multi-agent systems', *arXiv preprint arXiv:1303.4692* .

Anderson, E. F. (2003), 'Playing smart-artificial intelligence in computer games'.

Aschwanden, G., Halatsch, J. & Schmitt, G. (2008), Crowd simulation for urban planning, *in* 'Proceedings of eCAADe', Vol. 2008.

Baird, J. C. & Noma, E. J. (1978), *Fundamentals of scaling and psychophysics*, John Wiley & Sons.

Bard, E. G., Robertson, D. & Sorace, A. (1996), 'Magnitude estimation of linguistic acceptability', *Language* pp. 32–68.

Beltaief, O., El Hadouaj, S. & Ghedira, K. (2011), Multi-agent simulation model of pedestrians crowd based on psychological theories, *in* 'Logistics (LOGISTIQUA), 2011 4th International Conference on', IEEE, pp. 150–156.

Bevilacqua, F. (2013), 'Finite-state machines: Theory and implementation'. Last Accessed April 7, 2017.
URL: *http://gamedevelopment.tutsplus.com/tutorials/finite-state-machines-theory-and-implementation–gamedev-11867*

Bon, G. L. (1896), 'The crowd: A study of the popular mind'.

Borg, G. (1990), 'Psychophysical scaling with applications in physical work and the perception of exertion', *Scandinavian journal of work, environment & health* pp. 55–58.

Botea, A., Bouzy, B., Buro, M., Bauckhage, C. & Nau, D. (2013), Pathfinding in games, *in* 'Dagstuhl Follow-Ups', Vol. 6, Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik.

Botea, A., Müller, M. & Schaeffer, J. (2004), 'Near optimal hierarchical path-finding', *Journal of game development* **1**(1), 7–28.

Bouvier, E., Cohen, E. & Najman, L. (1997), 'From crowd simulation to airbag deployment: particle systems, a new paradigm of simulation', *Journal of Electronic imaging* **6**(1), 94–107.

Brosnan, A., Hamill, J., Dobbyn, S. & O'Sullivan, C. A. (2005), 'Animating humans on handlheld devices for interactive gaming'.

Browning, R. C., Baker, E. A., Herron, J. A. & Kram, R. (2006), 'Effects of obesity and sex on the energetic cost and preferred speed of walking', *Journal of Applied Physiology* **100**(2), 390–398.

Brunstrom, J. M., Shakeshaft, N. G. & Scott-Samuel, N. E. (2008), 'Measuring expected satietyin a range of common foods using a method of constant stimuli', *Appetite* **51**(3), 604–614.

CD Projekt RED (2015), 'The Witcher 3', Video Game.

Cornsweet, T. N. (1962), 'The staircase-method in psychophysics', *The American journal of psychology* **75**(3), 485–491.

Cui, X. & Shi, H. (2011), 'A*-based pathfinding in modern computer games', *International Journal of Computer Science and Network Security* **11**(1), 125–130.

de Carpentier, G. (2011), 'Boids: Simulating large flocks'. Last Accessed April 7, 2017.
URL: *http://www.decarpentier.nl/boids*

Dehghan, A., Idrees, H., Zamir, A. R. & Shah, M. (2014), Automatic detection and tracking of pedestrians in videos with various crowd densities, *in* 'Pedestrian and Evacuation Dynamics 2012', Springer, pp. 3–19.

Durupinar, F., Allbeck, J., Pelechano, N. & Badler, N. (2008), Creating crowd variation with the ocean personality model, *in* 'Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems-Volume 3', International Foundation for Autonomous Agents and Multiagent Systems, pp. 1217–1220.

Ehrenstein, W. H. & Ehrenstein, A. (1999), Psychophysical methods, *in* 'Modern techniques in neuroscience research', Springer, pp. 1211–1241.

Eklund, P. W., Kirkby, S. & Pollitt, S. (1996), A dynamic multi-source dijkstra's algorithm for vehicle routing, *in* 'Intelligent Information Systems, 1996., Australian and New Zealand Conference on', IEEE, pp. 329–333.

Ekman, G. (1959), 'Weber's law and related functions', *The Journal of Psychology* **47**(2), 343–352.

Emerson, P. L. (1986), 'Observations on maximum-likelihood and bayesian methods of forced-choice sequential threshold estimation', *Attention, Perception, & Psychophysics* **39**(2), 151–153.

Ennis, C., Peters, C. & O'Sullivan, C. (2011), 'Perceptual effects of scene context and viewpoint for virtual pedestrian crowds', *ACM Transactions on Applied Perception (TAP)* **8**(2), 10.

Fang, J., Qin, Z., Lu, Z. & Zhao, F. (2012), 'Fundamental diagram of pedestrian dynamics by safety interspace model', *arXiv preprint arXiv:1206.2127* .

Farell, B. & Pelli, D. G. (1999), 'Psychophysical methods, or how to measure a threshold and why', *Vision research: A practical guide to laboratory methods* **5**, 129–136.

Fechner, G. (1966), 'Elements of psychophysics. vol. i.'.

Ferber, J. (1999), *Multi-agent systems: an introduction to distributed artificial intelligence*, Vol. 1, Addison-Wesley Reading.

Freud, S. (1975), *Group psychology and the analysis of the ego*, number 770, WW Norton & Company.

Garcıa-Pérez, M. A. (1998), 'Forced-choice staircases with fixed step sizes: asymptotic and small-sample properties', *Vision research* **38**(12), 1861–1881.

Gescheider, G. A. (2013), *Psychophysics: the fundamentals*, Psychology Press.

Gill, A. et al. (1962), 'Introduction to the theory of finite-state machines'.

Go, J., Vu, T. D. & Kuffner, J. J. (2006), 'Autonomous behaviors for interactive vehicle animations', *Graphical Models* **68**(2), 90–112.

Golas, A., Narain, R., Curtis, S. & Lin, M. C. (2014), 'Hybrid long-range collision avoidance for crowd simulation', *IEEE transactions on visualization and computer graphics* **20**(7), 1022–1034.

Gold, J. I. & Shadlen, M. N. (2000), 'Representation of a perceptual decision in developing oculomotor commands', *Nature* **404**(6776), 390–394.

Gracely, R. H., Lota, L., Walter, D. & Dubner, R. (1988), 'A multiple random staircase method of psychophysical pain assessment', *Pain* **32**(1), 55–63.

Gröschel, A. (2011), 'Towards believable crowd simulation for interactive real-time applications', pp. 1–74.

Gu, Q., Yun, C. & Deng, Z. (2010), Perceiving motion transitions in pedestrian crowds, *in* 'Proceedings of the 17th ACM Symposium on Virtual Reality Software and Technology', ACM, pp. 199–202.

Guy, S. J., Curtis, S., Lin, M. C. & Manocha, D. (2012), 'Least-effort trajectories lead to emergent crowd behaviors', *Physical review E* **85**(1), 016110.

Guy, S. J., Kim, S., Lin, M. C. & Manocha, D. (2011), Simulating heterogeneous crowd behaviors using personality trait theory, *in* 'Proceedings of the 2011 ACM SIGGRAPH/Eurographics symposium on computer animation', ACM, pp. 43–52.

Hecht, S., Shlaer, S. & Pirenne, M. H. (1942), 'Energy, quanta, and vision', *The Journal of general physiology* **25**(6), 819–840.

Heïgeas, L., Luciani, A., Thollot, J. & Castagné, N. (2010), 'A physically-based particle model of emergent crowd behaviors', *arXiv preprint arXiv:1005.4405* .

Helbing, D. (2014), 'Social forces-revealing the causes of success or disaster (chapter 6 of digital society)', *Digital Society* .

Helbing, D. & Molnar, P. (1995), 'Social force model for pedestrian dynamics', *Physical review E* **51**(5), 4282.

INCONTROL (2016), 'Simulation as part of baggage handling operations'. Last Accessed April 7, 2017.
URL: *http://www.incontrolsim.com/simulation-part-baggage-handling-operations/*

Infinity Ward (2011), 'Call of duty: Modern warfare 3', Video Game.

IO Interactive (2016), ‘Hitman’, Video Game.

Kapadia, M., Shoulson, A., Boatright, C. D., Huang, P., Durupinar, F. & Badler, N. I. (2012), Whats next? the new era of autonomous virtual humans, *in* ‘Motion in Games’, Springer, pp. 170–181.

Kim, S., Guy, S. J., Manocha, D. & Lin, M. C. (2012), Interactive simulation of dynamic crowd behaviors using general adaptation syndrome theory, *in* ‘Proceedings of the ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games’, ACM, pp. 55–62.

Klein, S. A. (2001), ‘Measuring, estimating, and understanding the psychometric function: A commentary’, *Attention, Perception, & Psychophysics* **63**(8), 1421–1455.

Kok, V. J., Lim, M. K. & Chan, C. S. (2016), ‘Crowd behavior analysis: A review where physics meets biology’, *Neurocomputing* **177**, 342–362.

Korf, R. E. (1993), ‘Linear-space best-first search’, *Artificial Intelligence* **62**(1), 41–78.

Lance, F., Matheossian, D. & Poli, A. (2013), ‘City procedural modelling open source tool-kit’. Last Accessed April 7, 2017.
URL: *https://github.com/FlorianLance/City-procedural-modeling*

Le Bon, G. (2009), *Psychology of crowds*, Sparkling Books.

Leek, M. R. (2001), ‘Adaptive procedures in psychophysical research’, *Attention, Perception, & Psychophysics* **63**(8), 1279–1292.

Leggett, R. (2004), *Real-time crowd simulation: A review.*

Lemercier, S., Jelic, A., Kulpa, R., Hua, J., Fehrenbach, J., Degond, P., Appert-Rolland, C., Donikian, S. & Pettré, J. (2012), Realistic following behaviors for crowd simulation, *in* ‘Computer Graphics Forum’, Vol. 31, Wiley Online Library, pp. 489–498.

Levine, R. V. & Norenzayan, A. (1999), ‘The pace of life in 31 countries’, *Journal of cross-cultural psychology* **30**(2), 178–205.

Liu, Y., Liu, D., Badler, N. & Malkawi, A. (2011), Analysis of evacuation performance of merging points in stadiums based on crowd simulation, *in* ‘Proceedings of the 12th Conference of the International Building Performance Simulation Association (IBPSA, BS11)’, pp. 2651–2658.

Luo, L., Zhou, S., Cai, W., Low, M. Y. H. & Lees, M. (2009), Toward a generic framework for modeling human behaviors in crowd simulation, *in* 'Proceedings of the 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology-Volume 02', IEEE Computer Society, pp. 275–278.

McDonnell, R., Dobbyn, S., Collins, S. & O'Sullivan, C. (2006), Perceptual evaluation of lod clothing for virtual humans, *in* 'Proceedings of the 2006 ACM SIGGRAPH/Eurographics symposium on Computer animation', Eurographics Association, pp. 117–126.

McDonnell, R., Larkin, M., Dobbyn, S., Collins, S. & O'Sullivan, C. (2008), 'Clone attack! perception of crowd variety', *ACM Transactions on Graphics (TOG)* **27**(3), 26.

McDonnell, R., Newell, F. & O'Sullivan, C. (2007), Smooth movers: perceptually guided human motion simulation, *in* 'Proceedings of the 2007 ACM SIGGRAPH/Eurographics symposium on Computer animation', Eurographics Association, pp. 259–269.

McMahan, A. (2003), 'Immersion, engagement and presence', *The video game theory reader* **67**, 86.

Mehran, R., Oyama, A. & Shah, M. (2009), Abnormal crowd behavior detection using social force model, *in* 'Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on', IEEE, pp. 935–942.

Melo, M., Bessa, M., Debattista, K. & Chalmers, A. (2014), 'Evaluation of hdr video tone mapping for mobile devices', *Signal Processing: Image Communication* **29**(2), 247–256.

Nachmias, J. & Steinman, R. M. (1965), 'An experimental comparison of the method of limits and the double staircase-method', *The American journal of psychology* **78**(1), 112–115.

Nareyek, A. (2004), 'Ai in computer games', *Queue* **1**(10), 58.

O'Connor, S., Liarokapis, F. & Jayne, C. (2015), Perceived realism of crowd behaviour with social forces, *in* 'Information Visualisation (iV), 2015 19th International Conference on', IEEE, pp. 494–499.

O'Connor, S., Liarokapis, F. & Peters, C. (2013*a*), An initial study to assess the perceived realism of agent crowd behaviour in a virtual city, *in* 'Games and Virtual Worlds for

Serious Applications (VS-GAMES), 2013 5th International Conference on', IEEE, pp. 1–8.

O'Connor, S., Liarokapis, F. & Peters, C. (2013*b*), A perceptual study into the behaviour of autonomous agents within a virtual urban environment, *in* 'World of Wireless, Mobile and Multimedia Networks (WoWMoM), 2013 IEEE 14th International Symposium and Workshops on a', IEEE, pp. 1–6.

Ondřej, J., Pettré, J., Olivier, A.-H. & Donikian, S. (2010), A synthetic-vision based steering approach for crowd simulation, *in* 'ACM Transactions on Graphics (TOG)', Vol. 29, ACM, p. 123.

Paris, S., Pettré, J. & Donikian, S. (2007), Pedestrian reactive navigation for crowd simulation: a predictive approach, *in* 'Computer Graphics Forum', Vol. 26, Wiley Online Library, pp. 665–674.

Park, C.-S., Tahk, M.-J. & Bang, H. (2003), Multiple aerial vehicle formation using swarm intelligence, *in* 'AIAA Guidance, Navigation, and Control Conference and Exhibit', p. 5729.

Peirce, C. S. & Jastrow, J. (1884), 'On small differences in sensation'.

Pelechano, N., Allbeck, J. M. & Badler, N. I. (2007), Controlling individual agents in high-density crowd simulation, *in* 'Proceedings of the 2007 ACM SIGGRAPH/Eurographics Symposium on Computer Animation', SCA '07, Eurographics Association, Aire-la-Ville, Switzerland, Switzerland, pp. 99–108.

Pelechano, N. & Badler, N. I. (2006), 'Improving the realism of agent movement for high density crowd simulation', *University of Pennsylvania, Center for Human Modeling and Simulation* .

Peters, C., Dobbyn, S., Nameee, B. M. & O'Sullivan, C. (2003), 'Smart objects for attentive agents'.

Peters, C. & Ennis, C. (2009), 'Modeling groups of plausible virtual pedestrians', *IEEE Computer Graphics and Applications* (4), 54–63.

Peters, C., Ennis, C., McDonnell, R. & OSullivan, C. (2008), 'Crowds in context: Evaluating the perceptual plausibility of pedestrian orientations', *Proceedings of Eurographics Short Papers* pp. 227–230.

Portugal, R. & Svaiter, B. F. (2011), 'Weber-fechner law and the optimality of the logarithmic scale', *Minds and Machines* **21**(1), 73–81.

Poulton, E. C. (1968), 'The new psychophysics: six models for magnitude estimation.', *Psychological bulletin* **69**(1), 1.

Quinlan, J. R. (1986), 'Induction of decision trees', *Machine learning* **1**(1), 81–106.

Relkin, E. M. & Pelli, D. G. (1987), 'Probe tone thresholds in the auditory nerve measured by two-interval forced-choice procedures', *The Journal of the Acoustical Society of America* **82**(5), 1679–1691.

Reynolds, C. W. (1987), Flocks, herds and schools: A distributed behavioral model, *in* 'ACM Siggraph Computer Graphics', Vol. 21, ACM, pp. 25–34.

Reynolds, C. W. (1999), Steering behaviors for autonomous characters, *in* 'Game developers conference', Vol. 1999, pp. 763–782.

Roederer, J. G. (2008), *The physics and psychophysics of music: an introduction*, Springer Science & Business Media.

Schlauch, R. S. & Rose, R. M. (1990), 'Two-, three-, and four-interval forced-choice staircase procedures: Estimator bias and efficiency', *The Journal of the Acoustical Society of America* **88**(2), 732–740.

Sekuler, R., Watamaniuk, S. N. & Blake, R. (2002), 'Perception of visual motion', *Stevens Handbook of Experimental Psychology. Third edition. H. Pashler, series editor. S. Yantis, volume editor. J. Wiley Publishers. New York* .

Shao, W. & Terzopoulos, D. (2005), Autonomous pedestrians, *in* 'Proceedings of the 2005 ACM SIGGRAPH/Eurographics symposium on Computer animation', ACM, pp. 19–28.

Shelton, B. & Scarrow, I. (1984), 'Two-alternative versus three-alternative procedures for threshold estimation', *Attention, Perception, & Psychophysics* **35**(4), 385–392.

Shining Rock Software (2014), 'Banished', Video Game.

Simpson, W. A. (1988), 'The method of constant stimuli is efficient', *Attention, Perception, & Psychophysics* **44**(5), 433–436.

Song, Y., Gong, J., Li, Y., Cui, T., Fang, L. & Cao, W. (2013), 'Crowd evacuation simulation for bioterrorism in micro-spatial environments based on virtual geographic environments', *Safety science* **53**, 105–113.

Stout, B. (1996), 'Smart moves: Intelligent pathfinding', *Game developer magazine* **10**, 28–35.

Szymanezyk, O., Duckett, T. & Dickinson, P. (2012), Agent-based crowd simulation in airports using games technology, *in* 'Transactions on Computational Collective Intelligence VIII', Springer, pp. 192–213.

Tung, B. & Kleinrock, L. (1996), 'Using finite state automata to produce self-optimization and self-control', *IEEE transactions on parallel and distributed systems* **7**(4), 439–448.

Ubisoft Quebec (2015), 'Assassin's Creed Syndicate', Video Game.

Ulicny, B. & Thalmann, D. (2001), Crowd simulation for interactive virtual environments and vr training systems, *in* 'Computer Animation and Simulation 2001', Springer, pp. 163–170.

Ulicny, B. & Thalmann, D. (2002), Crowd simulation for virtual heritage, *in* 'Proc. First International Workshop on 3D Virtual Heritage', number VRLAB-CONF-2007-042, pp. 28–32.

Ulrich, R. & Miller, J. (2004), 'Threshold estimation in two-alternative forced-choice (2afc) tasks: The spearman-kärber method', *Perception & Psychophysics* **66**(3), 517–533.

University of Greenwich (2012), 'Scientists launch world's most advanced crowd simulation and evacuation software'. Last Accessed April 7, 2017.
URL: *http://www2.gre.ac.uk/about/news/articles/2012/a2140-ed-galea-exodus-software*

Watson, A. B. & Pelli, D. G. (1983), 'Quest: A bayesian adaptive psychometric method', *Perception & psychophysics* **33**(2), 113–120.

Weber, E. H. (1978), *EH Weber: The sense of touch*, Academic Pr.

Wetherill, G. & Levitt, H. (1965), 'Sequential estimation of points on a psychometric function', *British Journal of Mathematical and Statistical Psychology* **18**(1), 1–10.

Wier, C. C., Jesteadt, W. & Green, D. M. (1976), 'A comparison of method-of-adjustment and forced-choice procedures in frequency discrimination', *Attention, Perception, & Psychophysics* **19**(1), 75–79.

Yarnitsky, D., Sprecher, E., Zaslansky, R. & Hemli, J. A. (1995), 'Heat pain thresholds: normative data and repeatability', *Pain* **60**(3), 329–332.

Yegnanarayana, B. (2009), *Artificial neural networks*, PHI Learning Pvt. Ltd.

Zhou, S., Chen, D., Cai, W., Luo, L., Low, M. Y. H., Tian, F., Tay, V. S.-H., Ong, D. W. S. & Hamilton, B. D. (2010), 'Crowd modeling and simulation technologies', *ACM Transactions on Modeling and Computer Simulation (TOMACS)* **20**(4), 20.

Zhou, S., Sun, Y., Lu, L. & Chen, Z. (2006), Fire simulation model based on particle system and its application in virtual reality, *in* '16th International Conference on Artificial Reality and Telexistence–Workshops (ICAT'06)', IEEE, pp. 642–645.

Zhu, K. & Jiang, M. (2010), Quantum artificial fish swarm algorithm, *in* 'Intelligent Control and Automation (WCICA), 2010 8th World Congress on', IEEE, pp. 1–5.